

Competing Moral Minds?

Reexamining Moral Conflict between the Left and Right in American Politics

Abstract

What drives ideological and partisan divisions in contemporary American politics? An influential line of research suggests they are rooted in disagreement about the nature of morality. While the left uses principles of care and fairness in moral judgment, the right considers loyalty, authority, and sanctity to be additional morally relevant values. This creates a “moral empathy gap” that makes it difficult for people to understand the perspective of their political opponents, and fosters intolerance and gridlock. We argue that evidence for a sharp moral divide separating the left and right rests largely on a survey measure with significant limitations—the moral foundations questionnaire. We outline the methodological issues associated with this measure and use two alternative strategies to examine moral conflict between the American left and right. Across six distinct empirical tests with U.S. adults, we find that moral differences between these political groups are relatively small.

What drives ideological and partisan divisions in contemporary American politics? An influential line of research using *moral foundations theory* (hereafter MFT; Graham et al. 2013, Haidt 2012) suggests they are rooted in fundamental disagreements about the nature of morality. According to MFT, there are at least five distinct concerns that form the bases for human moral judgment and decision making: care, fairness, authority, loyalty, and sanctity. Care and fairness are “individualizing” foundations because they promote individual welfare and rights. The latter three are “binding” foundations because they maintain group cohesiveness and social order. A growing body of research suggests that while the left primarily uses care and fairness in moral judgment, the right also considers loyalty, authority, and sanctity to be legitimate moral concerns (see Kivikangas et al. 2021 for a review and meta-analysis; e.g., Graham, Haidt, and Nosek 2009, Haidt and Graham 2007). Such disagreement creates a “moral empathy gap” and predictable divisions between the left and right over political values (Federico et al. 2013), policy preferences (Kertzer et al. 2014), and candidate evaluation (Weinschenk and Dawes 2019). The present paper reconsiders this claim. We argue that existing evidence for a sharp moral divide between the left and right rests largely on a survey measure with significant limitations—the moral foundations questionnaire (hereafter MFQ). We outline the methodological issues associated with this measure. We then employ two novel strategies for testing the relationship between moral foundations and political orientation. Across six distinct empirical tests, and multiple measures of political preferences, we find that moral differences between the American left and right are relatively small.

Our work makes an important contribution to an emerging literature which critically evaluates the role of moral conflict in American politics. Smith et al. (2017) analyze panel data and find that changes in moral foundations (measured using the MFQ) fail to predict changes in

ideology. Along similar lines, Hatemi, Crabtree, and Smith (2019) suggest that ideological identity is a causal *antecedent* to moral foundations as measured by the MFQ. Similarly, Ciuk (2018) finds that political group cues influence citizens' responses to the MFQ. This line of work suggests that moral disagreements in the mass public may be partly endogenous to political identification rather than a primary cause of political sorting. While this body of research raises doubts about the causal primacy of moral foundations, it is entirely based on responses to the MFQ which, as we outline below, has significant problems. The present research takes a different approach. We explore the primary observable implication of differences between the moral preferences of the left and right—that they make different moral choices—and find that differences between the left and right, of whatever origin, are small in magnitude.

These results have important implications for understanding polarization and mass political behavior in contemporary American politics. Indeed, some argue that there is a “moral empathy gap” between the left and right, such that people find it difficult to see the world from the perspective of their political opponents (Ditto and Koleva 2011), creating political intolerance and gridlock. As Graham et al. (2013, 75-76) suggest, this “could help to explain many aspects of the culture war... the intractability of the debates, and the inability of the two sides to... understand each other (because their moral visions were based on... differences in the very foundations upon which moral arguments could rest).” It matters a great deal if political conflict is rooted in, or expressed as, disagreement over the *principles* that define right and wrong rather than disagreement about *how to apply* common moral concerns within a given context. The former suggests pessimism about the prospects for reasoned debate on many issues in American politics because the two sides cannot agree on the terms upon which such a discussion would be based. Our results, however, suggest the possibility for reasoned persuasion

and the uncovering of common moral ground between opposing sides. At minimum, our work suggests the need to look to other sources to understand the seeming intractability of contemporary American politics.

Moral Foundations Theory and the Moral Foundations Questionnaire

MFT argues that moral systems are “interlocking sets of values, practices, institutions, and evolved psychological mechanisms that work together to...regulate selfishness and make social life possible” (Haidt 2008, 70). Due to recurring challenges faced by our evolutionary forebears, human cognition is “organized in advance of experience” to learn certain types of moral principles rather than others. For example, just as it’s easier to teach a child to be afraid of a spider than a flower, it is easier to teach in-group loyalty than cosmopolitanism. Further, because there were diverse adaptive challenges associated with social life for human ancestors, there are multiple concerns that guide human moral judgment and decision making. MFT proposes five foundations, defined by both a positive and negative concept, and associated with five recurring selection pressures (described in parentheses): care/harm (caring for children), fairness/cheating (reaping benefits of dyadic interactions), loyalty/betrayal (forming cohesive coalitions), authority/subversion (navigating social hierarchies), and sanctity/degradation (avoiding communicable diseases). These five are (often) further grouped according to locus of concern with the first two referred to as “individualizing” foundations, due to their concern for the well-being of individuals, and the last three referred to as “binding” foundations due to their concern with individuals’ behavior in service of the group.

This framework has spurred innovative and influential research across a variety of domains, including consumer choice (Watkins, Aitken, and Mather 2016), charitable donations (Winterich, Zhang, and Mittal 2012), impression formation (Clifford and Jerit 2013, Smith et al.

2019), political belief systems (Federico et al. 2013), and political participation (Johnson et al. 2014). Most importantly for our purposes, research suggests that the five foundations are among the strongest predictors of political ideology (Graham, Haidt, and Nosek 2009) and are important to policy-relevant beliefs on issues such as foreign affairs (Kertzer et al. 2014), stem-cell research (Clifford and Jerit 2013), the environment (Feinberg and Willer 2013), and the culture wars (Koleva et al. 2012). In this sense, moral foundations theory has demonstrated its “pragmatic validity” by providing novel insights into important topics, including contemporary American politics (Graham et al. 2013, 71) and the left-right conflict that defines it (Haidt 2012).

Importantly, however, empirical evidence for these claims rests largely on a single research strategy—the MFQ.¹ The MFQ asks respondents to rate the relevance of different moral concerns to their judgments of right and wrong (Form A) and also about their agreement with statements that reflect the importance of such concerns (Form B). We list all 30 MFQ items and their associated foundations in the Supplemental Appendix. While the MFQ has inarguably made a valuable contribution to the literature on mass politics, it has significant limitations. We review these issues in depth before turning to our own approach.

¹ To be sure, some research uses other methods (see Graham et al. 2013 for a review). Studies that employ psychophysiological measures, as well those that analyze political discourse, both find support for some of these claims (Neiman et al. 2016, Sterling and Jost 2018, but see Frimer 2020). But the bulk of empirical support is from studies using the MFQ.

First, as shown in the Supplemental Appendix, the MFQ has concerning measurement properties. Using three distinct national samples², we show that many of the MFQ items load poorly on their respective latent factors. With the exception of the sanctity foundation, we find a clear pattern to these loadings: items using Form B (agree-disagree format) have consistently smaller loadings than those using Form A (relevance judgments). This suggests measurement variance may be quite important to the MFQ. Consistent with this hypothesis, model fit improves substantially when we account for individual differences in the use of the two forms—that is, differing response tendencies unrelated to the substantive content of the items. This is reflected in the standardized loadings of the items on latent factors for each form (i.e., “methods” factors). In every case they are non-negligible, and in many cases—especially for Form A—variance in items attributable to method exceeds that attributable to moral content. Ultimately, many items are better measures of individual differences in respondents’ use of the survey instrument than individual differences in moral priorities.

Second, recent research finds only two factors in exploratory factor analyses of the MFQ (Smith et al. 2017). In confirmatory analyses, research finds only a modest increase in fit from a two-factor model—specifying only individualizing and binding foundations—to a model that includes all five moral foundations as latent variables (Graham et al. 2011; Iurino and Saucier 2018; Nilsson and Erlandsson 2015). We find further support for a two-factor model in the covariance structure analyses of the MFQ in the three datasets reported in the Supplemental Appendix. Foundations that share a superordinate domain (i.e., individualizing or binding) are

² The 2012 Cooperative Congressional Election Study, the 10th wave of The American Panel Study, and the fourth wave of the National Survey of Youth and Religion. To our knowledge, these are the first covariance structure analyses of the MFQ using these datasets.

correlated at near-unity, and the covariance matrices of the latent factors are not always positive-definite. This is consistent with other recent work (Iurino and Saucier 2018), and suggests the five-factor model is misspecified. Moreover, in one of our datasets we find that the *interdomain* correlations among the foundations (e.g., loyalty and care) are often strongly and negatively correlated with one another. Overall, past work, as well as our own analyses, suggest the MFQ does not strongly differentiate moral content within the individualizing and binding domains. This is a significant limitation. As we will show later in this paper, moral priorities revealed through choices do not cleanly separate moral foundations into individualizing and binding categories.

The MFQ has at least two other significant limitations. First, respondents consider moral concerns in isolation, but moral decision making typically requires tradeoffs among conflicting considerations, such as aggregate welfare versus individual rights or obedience to authority versus preventing harm. To borrow language from research on human values, moral considerations may work in a comparative and competitive fashion (Schwartz 1992). As such, the MFQ may elicit exaggerated importance ratings for some foundations if people do not spontaneously bring to mind the fact that other moral considerations are likely to be relevant in most cases of actual moral judgment.³ A similar problem has been found in the voting literature, where citizen ratings of the personal importance of issues, considered in isolation, have little relationship to the revealed importance of issues in vote choice (Leeper and Robinson 2018).

³ Though some work has considered the importance of ranking moral foundations in empirical research (Feinberg and Willer 2015; Monroe, Wyngaarden III and Plant 2021; Waytz, Dungan and Young 2013), these scholars consider only a small number of moral foundations at a time.

Second, left-right differences in binding foundation importance may be partly “baked-in” to the MFQ. Specifically, many of the items used to measure the binding foundations evoke stereotypically right-wing concerns. For example, sanctity is measured with items tapping attitudes toward religion and chastity; loyalty is measured with attitudes toward patriotism; and authority is measured with attitudes toward gender roles and societal traditions. Yet, as Schein and Gray (2018, 55) argue, on some issues it is the left that emphasizes collectivistic and purity-related concerns (e.g., on the environment). The lack of items allowing for such left-wing interpretations of the binding foundations makes it “unclear whether measures such as the MFQ reveal new knowledge about individual differences in morality or simply repackage a subset of well-known political differences with pithy names.”

This possibility finds support in recent research on the causal ordering of the MFQ in relation to political attitudes and ideology. As stated above, Hatemi, Crabtree and Smith (2019) and Smith et al. (2017), using panel data, find that political ideology is more stable than the MFQ within individuals over time and that lagged values of the MFQ are largely unrelated to changes in ideology. Moreover, Hatemi et al. (2019) find that lagged values of ideology predict changes in the MFQ, providing evidence that cross-sectional correlations of these variables may be due, in part, to the causal influence of ideology on the MFQ. Consistent with this observational work, Ciuk (2018) provides experimental evidence that political group cues (particularly out-group cues) shape responses to the MFQ.

Taken together, recent research suggests that the MFQ has important methodological limitations and that cross-sectional correlations of the MFQ with political preferences may be due, in part, to the influence of the latter on the former. In the remainder of the article, we reconsider the relationship between moral priorities and political preferences in American

politics using two different strategies designed to capture respondents' moral priorities. To be clear, our purpose is not to adjudicate debates about the causal ordering of moral foundations and political preferences. Rather, our concern is whether sharp moral differences between the left and right replicate when using alternative measures of moral priorities beyond the MFQ. If these differences are idiosyncratic to the MFQ, the issue of causal ordering may be moot. On the other hand, if such differences are robust to measurement strategy, further exploration of causal ordering with alternative measures is warranted.

Our first measurement approach asks respondents to explicitly rank-order the importance of the five moral foundations, while our second approach elicits the relative importance of the foundations indirectly by asking respondents to make choices about the relative moral wrongness of acts. In both cases, respondents are forced to consider, and choose, what moral considerations are most important to them. While such instruments are new to MFT research, they are used frequently in research on related topics. Scholarship on core political values, for example, argues that values do not work in isolation. Rather, they work in a comparative and competitive fashion such that the *relative* importance of values guides behavior and choice (Schwartz 1992), and insofar that moral foundations work in the same general way as values (see Graham et al. 2009, 1030), arguments pertaining to the measurement of values ought to be applied to MFT.

To summarize, where previous critiques of MFT in political science focus mainly on the causal relationships between political identities and the MFQ (Ciuk 2018; Hatemi et al. 2019; Smith et al. 2017), we focus on replicating prior results of sharp moral divisions using two distinct measurement approaches. Ultimately, our results run contrary to previous work using the MFQ: there are more similarities than differences between liberals and conservatives in their

moral priorities. We conclude by considering the broader implications of our work and how it contributes to the existing literature. We consider each measurement approach in turn.

Study 1

Data

We collected three survey datasets for Study 1. Four-hundred and six respondents were recruited through Amazon.com's Mechanical Turk interface in June 2017; 1,478 respondents were recruited through Dynata (formerly SSI/Research Now) in April 2019; and 1,777 respondents were collected through Lucid in March and April of 2022.⁴

Moral Foundation Ranking Task

In each survey, respondents completed an importance ranking task using the method of triads (see Ciuk and Jacoby 2015, 713-714). They were first shown definitions of the five moral foundations. They were then shown a series of 10 screens, each of which displayed a set of three of the five foundations, and were asked to indicate their most and least important foundation in each set. The moral foundation definitions were identical for the MTurk and Dynata samples. In the Lucid sample, half of respondents (N=888) were randomly assigned to receive the same definitions as the MTurk and Dynata samples, and half (N=889) were randomly assigned to receive a new set of definitions (all definitions and additional details are in the Supplemental Appendix).

⁴ These sample sizes include all respondents who completed the survey. In the Lucid survey, respondents were removed if they failed any attention checks, were associated with a duplicate respondent ID, or were flagged by Qualtrics as having a high probability of being a bot or being a fraudulent data point. More information is available in the Supplemental Appendix.

Measurement of Political Orientation

We measure ideological identification and partisanship in all three samples using standard seven-point scales. For each measure separately, we divide the sample into three groups, consisting of left-wing (Democrats/Liberals, including leaners), moderate (Independents/Moderates), and right-wing (Republicans/Conservatives, including leaners) citizens. The Dynata and Lucid data samples (but not MTurk) also contain several policy items on various salient issues in American politics which are used to construct three issue-based measures of left-right preferences: a general left-right issue dimension, an economic dimension, and a social/cultural dimension. For each of these three measures, we divide respondents into three groups based on tertiles of these scales (left, moderate, and right). Further details and item wordings are in the Supplemental Appendix.

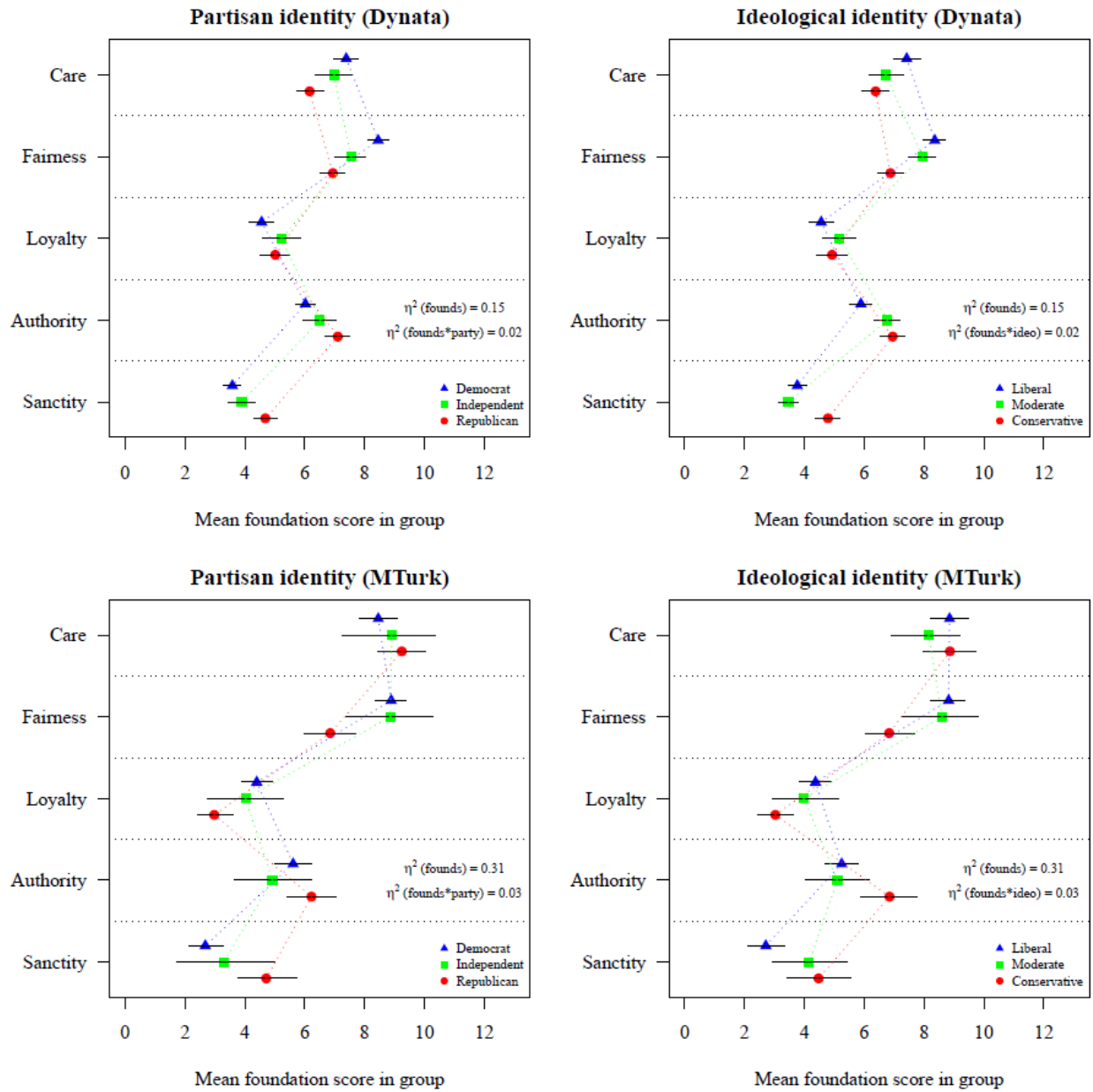
With these data, our goals are to describe the typical pattern of moral priorities for left-wing, moderate, and right-wing citizens; compare these priorities across groups; and assess the extent to which variation in moral priorities varies both across and within groups. We keep the analysis simple so that inferences do not rest on (potentially contentious) modeling choices. For each respondent in each sample, we calculate the number of times a given foundation was ranked above another foundation across the ten ranking tasks (i.e., the top-ranked foundation in each triad gets a rank-score of 2 because it was ranked above two other foundations, the second-ranked gets a rank-score of 1, and the bottom-ranked gets a rank-score of 0). Since each foundation appears in a total of six tasks, the maximum possible rank-score is 12. We then calculate the mean score for each of the five foundations within each political group (e.g., separately for Democrats, Independents, and Republicans). Since there are dependencies in the rank-scores within each respondents' set of five scores, we bootstrap the standard errors for these

means. Specifically, we took 1,000 random samples of respondents, with replacement, in each sample separately. For each of these sets of 1,000 samples, we calculated the relevant means. We then report 2.5th and 97.5th percentiles of these sample means as 95% confidence intervals.

In addition to the sample means, we also calculate a variance decomposition to assess the strength of the relationship between political preferences and moral priorities. Specifically, we calculate the percentage of the variation (η^2) in individuals' ranking scores that is explained by (1) the categorical foundation factor (i.e., the explanatory power of knowing which foundation a score is associated with) and (2) by the *interaction* of political preferences (left-wing, moderate, or right-wing) with the foundation factor. The latter is the additional variation explained in ranking scores when we calculate separate means for each political group, relative to means for the entire sample combined. The remaining variance—unaccounted for by (1) or (2)—is that attributable to individual variation in moral priorities around the group-specific means for the five foundations. It is thus variation that is idiosyncratic to individuals.

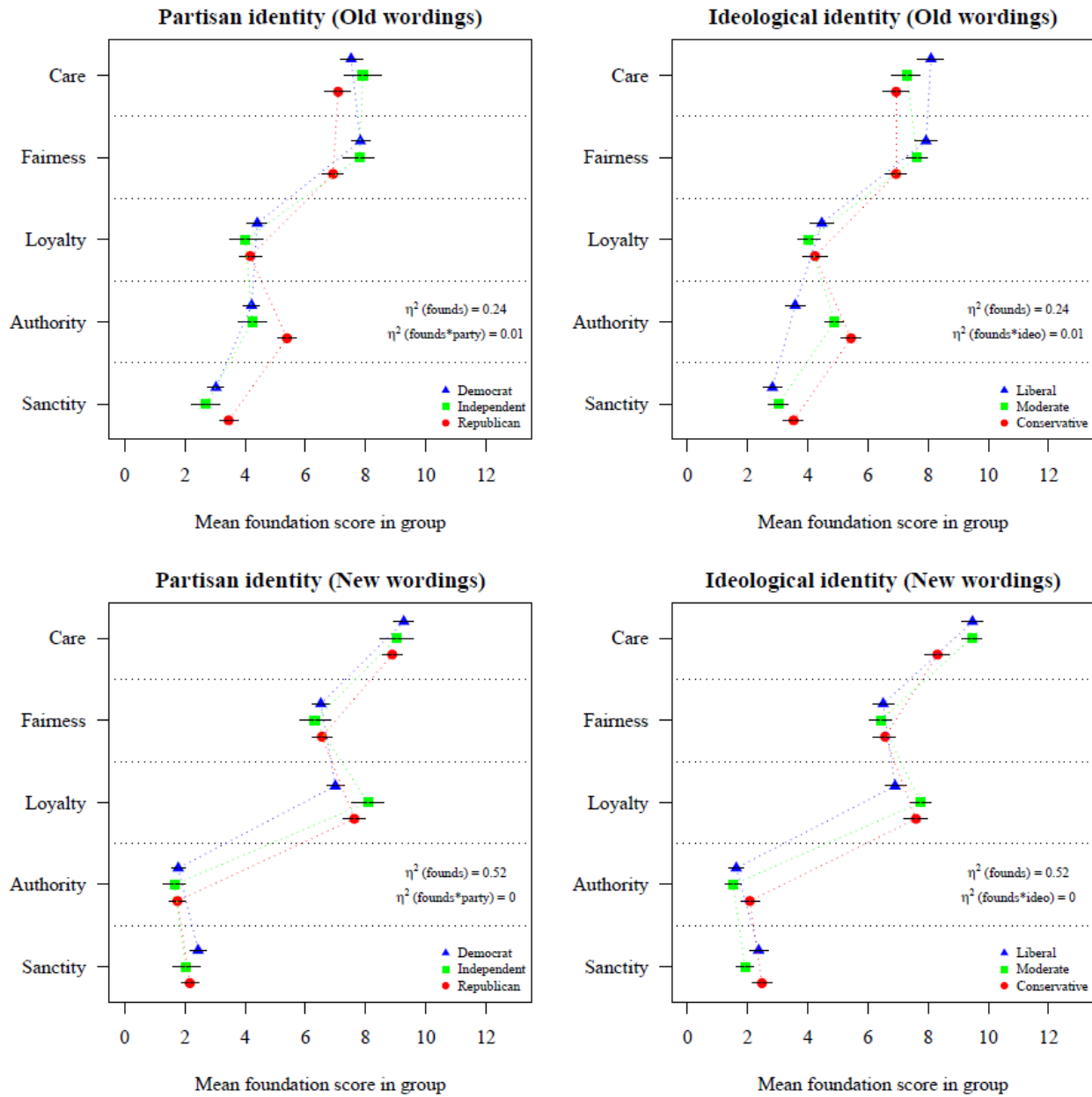
For purposes of interpretation, the variation explained by the foundations variable alone tells us how much consensus there is in the population as a whole. The closer this value gets to 1, the more agreement there is about the relative importance of these moral foundations in our overall samples. The variation explained by the interaction of the foundations factor with political orientation tells us how much moral conflict exists across political groups in our sample. At the extreme, when this value is 1, it implies that members of each political group have identical moral priorities, and all variation is due to political group differences.

Figure 1. Study 1 Estimates for Dynata and MTurk Surveys



Notes: Each point represents the mean rank-score for a particular foundation in a particular group. Extended lines are bootstrapped 95% confidence bounds. $\eta^2(\text{founds})$ is the variance in rank-scores explained by the five foundations, and $\eta^2(\text{founds*group})$ is the additional variance explained by group differences in the pattern of foundation rank-scores.

Figure 2. Study 1 Estimates for Lucid Survey (Old versus New Foundation Definitions)



Notes: Each point represents the mean rank-score for a particular foundation in a particular group. Extended lines are bootstrapped 95% confidence bounds. $\eta^2(\text{foundcs})$ is the variance in rank-scores explained by the five foundations, and $\eta^2(\text{foundcs*group})$ is the additional variance explained by group differences in the pattern of foundation rank-scores.

Results

We report our results for Study 1 in Figures 1 and 2. Figure 1 shows the results for the MTurk and Dynata samples, and Figure 2 shows the results for the Lucid sample (“Old” and “New” wordings separately – see Supplemental Appendix). We restrict our discussion to the results for partisan and ideological identity, as these have been the focus of most discussion in the literature. We report the results for the remaining three political orientation variables (based on issues) in the Supplemental Appendix. They are very similar and offer the same substantive conclusions.

Across all four samples, and across the two distinct sets of foundation definitions, we find a very similar result. A substantial proportion of the variation in ranking scores is explained by differences across the five foundations themselves: 15%, 31%, 24%, and 52%, for the Dynata, MTurk, Lucid “old” definitions, and Lucid “new” definitions, respectively. Using the “old” definitions, care and fairness tend to receive the highest rank-scores, followed by authority, with loyalty and sanctity receiving the lowest. In the Dynata data, authority receives scores that are comparable to care and fairness, while in the MTurk and Lucid samples, it scores quite a bit lower, and is comparable to loyalty and sanctity. Using the “new” definitions, care receives the highest rank-scores, followed by loyalty and fairness in a close second and third, with authority and sanctity scoring much lower. Interestingly, the new definitions shift the importance of loyalty upwards at the expense of both authority and sanctity, and there is a clear separation between the average importance of the latter two foundations and the remaining three. Importantly, the difference in results comparing the “old” definitions to the “new” ones in the Lucid sample can be interpreted as a causal effect of the wordings, because respondents were

randomly assigned to these two tasks within the same sample. We thus demonstrate the importance of how these foundations are described to their perceived importance.

In contrast, little additional variance is attributable to differences in foundation means across political groups: 2%, 3%, 1%, and <1%, for the Dynata, MTurk, Lucid “old” definitions, and Lucid “new” definitions, respectively. These values are identical for both partisanship and ideology. To be sure, the variation that exists is roughly consistent with what is expected based on past work: the individualizing foundations typically receive higher ranking scores among left-wing respondents, while the binding foundations typically receive higher scores among right-wing respondents. However, these differences are often small. Moreover, in several cases, the patterns across the five foundations do not clearly map to the individualizing versus binding foundation distinction. For example, in the Lucid “old” definitions sample, the left and right only diverge substantially, and in the expected direction, on the importance of authority, and the left actually scores slightly higher on the importance of loyalty. In the Lucid “new” definitions sample, there is a clear separation, for all political groups, between care, fairness, and loyalty on one side, and authority and sanctity on the other.

This also reinforces the limitations of the MFQ with respect to its tendency to collapse to a two-dimensional—or even one-dimensional—structure. Both the Dynata and the MTurk surveys included 10-item measures of the MFQ. In the Supplemental Appendix, we show that the MFQ has a distinct pattern of results across measures of political preferences. First, the overall pattern in means, across foundations, comparing groups of the political left and right, largely conforms to past findings in the literature. Specifically, left-wing groups place greater importance on care and fairness relative to loyalty, authority, and sanctity, while right-wing groups tend to place roughly equal importance on both individualizing and binding foundations,

or place greater importance on the binding foundations. Moreover, left-wing groups tend to place greater importance on the individualizing foundations relative to right-wing groups, while right-wing groups place greater importance on the binding foundations relative to left-wing groups, though this is less consistent for the loyalty foundation. In each case, the variance explained by foundations alone is much less than with our alternative measurement strategy, indicating less agreement over foundation importance in the population as a whole. With one exception, the addition of the interaction of foundations with political preferences explains more variance than the comparable estimates using our alternative measurement approach. In the case of the MTurk data using the MFQ, the difference is quite large: the addition of party and ideology interactions explains an additional 8% and 11% of the variance above including the foundations alone. Given that we only have the 10-item measure of the MFQ available in these data, these estimates are likely to underestimate the differences across groups in MFQ due to measurement error.

Study 2

Data

Our second study avoids self-reported rankings altogether. Here, we *elicit* moral priorities from decisions about the relative immorality of specific acts. We collected two datasets, both which are non-probability samples collected using Lucid's Fulcrum Marketplace (Coppock and McClellan 2019; see Supplemental Appendix). The first dataset was collected in August of 2018 and consists of 994 completes. The second was collected in March and April of 2022 and consists of 916 completes.⁵

⁵ Completes are sample sizes after attention checks and exclusions. See footnote 4 and the Supplemental Appendix for details.

Moral Choice Task

In both studies, respondents first completed a set of demographic items along with questions assessing political attitudes and identities. Then, each respondent was presented with a series of binary choice tasks. Each option in each task consisted of a one-sentence scenario describing an act. Respondents were asked to choose the act that is “more morally wrong.”

The acts for the 2018 survey were chosen from the set of Moral Foundations Vignettes (MFV, Clifford, Iyengar, et al. 2015), with each act designed to depict a behavior that violates only a single moral foundation, and with all acts roughly normed to be equal in severity. We selected five acts (with roughly equivalent average wrongness ratings) from each moral foundation category in the MFV database, which makes a distinction between physical and emotional harm, and includes “liberty/oppression” as a sixth foundation. The MFV database also includes a category of amoral violations of social norms, or matters of taste or personal preference, and we selected five of these acts as well. We list all acts used in this study in the Supplemental Appendix. The advantage of the MFV is that it was created specifically for the purpose of isolating individual moral foundations and maintaining roughly equivalent levels of moral wrongness. This necessarily limits generalizability, however. First, acts that violate only a single moral foundation may be the exception rather than the rule. Second, equating severity across acts may lead to less representative acts in some domains relative to others. For example, it may be the case that the prototypical violation of the care/harm foundation is much more severe than the prototypical violation of the authority/subversion foundation.

For these reasons, the acts for the 2022 survey were assembled by the researchers to span a wider range of violations within each foundation (also listed in the Supplemental Appendix). For example, acts for the care/harm foundation include both making fun of someone’s clothes as

well as killing someone while drunk driving. The acts for the sanctity/degradation foundation include both eating with your mouth open in front of others and cursing God. The set of acts thus has significant variation in both content and severity. Importantly, while we made sure to include an equal number of acts we believed, a priori, would be seen to violate the five moral foundations, we did not restrict our set to those that violate *only* a single foundation. That is, by design, and in contrast to the MFV, we cannot identify each of our acts with only a single foundation in the subsequent analysis.

To overcome this issue, we collected a second dataset of 340 respondents (after exclusions; see Supplemental Appendix) using Amazon.com's Mechanical Turk. Each respondent was randomly assigned to one of the five primary moral foundations and received a definition of that foundation (see Supplemental Appendix). They were then presented with a series of randomly drawn acts from our assembled set and asked "to rate the extent to which each act violates the principle of [FOUNDATION]" on a five-point scale from "not a violation" to "an extreme violation." Each respondent separately rated a total of 30 distinct acts for their assigned foundation. Across all respondents, each act thus receives multiple violation ratings for each of the five foundations. As we describe further below, we use these ratings to elicit the relative importance of each foundation in the choices of our Lucid respondents.

Analysis Strategy

In the 2018 survey, respondents made 20 choices between two randomly generated acts with the only restriction on randomization being that the two acts presented in each task come from different MFV categories. We follow Hainmueller, Hopkins, and Yamamoto (2014) regarding the analysis of conjoint studies. Specifically, each respondent is represented in the data once for each act in each choice task, for a total of forty observations per person. The dependent

variable is coded one if the respondent chose that act as more morally wrong and zero otherwise. We regress this indicator on dummy indicators for the eight act groups (seven moral foundations—two for care with the additional liberty foundation—along with the amoral category). For ease of interpretation, we include the amoral dummy as a predictor and fix the intercept to zero, thus estimating the marginal means for each group of acts. Marginal means are interpreted as the estimated probability that a randomly drawn respondent will choose a particular act, given that it is a member of a particular moral foundation category, averaging over the distribution of all possible alternatives. We estimate all models by ordinary least squares with clustered standard errors by respondent.

In the 2022 survey, respondents made 30 choices between independently randomly selected pairs of acts. We regress the choice between the two acts (“A” or “B”) on a set of five predictors, each of which represents the difference in the extent to which the two acts violate a particular moral foundation, as rated by our MTurk respondents. For example, the first predictor subtracts the average violation severity score for act B on the care/harm foundation from the violation severity score for act A on that same foundation. Thus, all predictors are expected to be associated with choosing act A as more morally wrong, because higher scores indicate more severe violations of the moral foundations. Our interest is in the relative magnitude of the estimated coefficients across the five foundations and across various political groups of interest. The larger the coefficient for a given foundation, the more important that foundation is in predicting the moral judgments of a particular group.

Our moral violation severity scores are based on the average judgments from the MTurk respondents described above. For each act, we average, separately for each moral foundation, all available judgments about how much that act violates the respective foundation. Importantly,

these ratings are only a sample, and thus there is uncertainty in the average rating. To propagate this uncertainty into the estimation of coefficients in the choice regression, we bootstrap all estimates. Specifically, we took 1,000 random samples, with replacement, from both the set of MTurk respondents and the set of Lucid respondents. For each of these 1,000 samples, we calculated the average violation scores for all acts for all five foundations. We then regressed the choices of the sampled Lucid respondents on the severity differences scores calculated from the sampled MTurk respondents and saved the estimated coefficients.⁶ We use the variation in the estimated coefficients to calculate 95% confidence bounds for all coefficients. All models are estimated with ordinary least squares.

In both surveys, we run these models separately for self-identified Republicans, Independents, Democrats, Liberals, Moderates, and Conservatives. As we did in Study 1, we also construct three issue-based measures of political orientation (see Supplemental Appendix).

Results

Figure 3 displays the results for the two surveys with top two panels showing results for the MFV acts (2018 survey) and the bottom two panels the results for the newly assembled acts (2022 survey). As in Study 1, we focus our discussion on partisanship and ideology. We provide the results for the three other dimensions in the Supplemental Appendix (the results are similar).

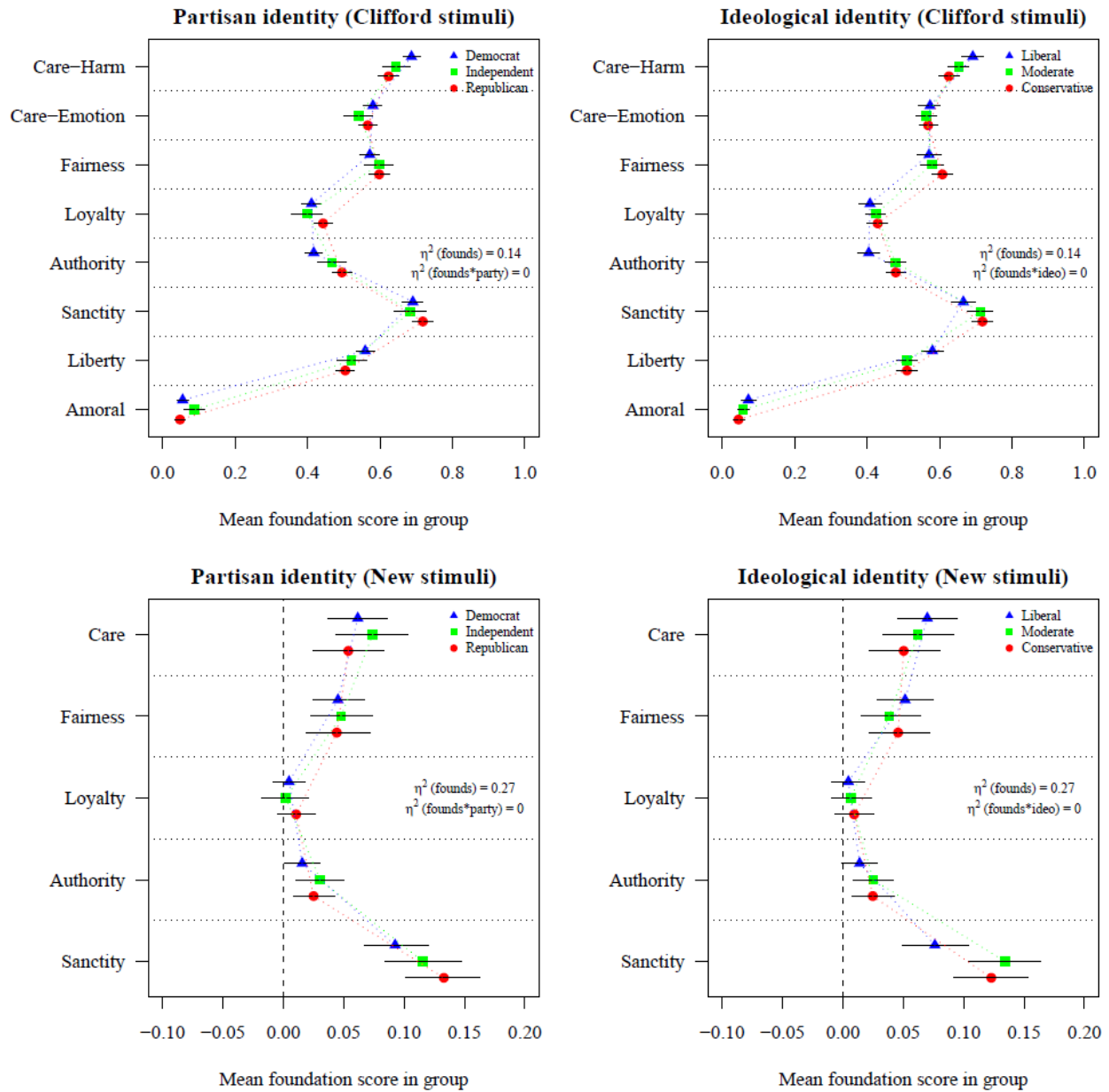
Beginning with the MFV, the results again suggest that little additional variance is explained by the interaction of political partisanship and ideology with moral foundations. Specifically, knowing the foundation with which an act is associated explains about 14% of the variance in the choices respondents make, while the interaction of the foundation indicators with

⁶ Sampling at the respondent level, rather than the judgment or choice level, ensure that uncertainty bounds reflect clustering within respondents.

partisanship and ideology explains almost none. This can be seen in the plot of the marginal means, which show meaningful variation across the eight moral categories, but minimal variation across political groups within each of these categories. Both care foundations, along with fairness and sanctity, have the strongest relationships with choice behavior. For example, averaging over all possible alternatives, an act associated with the sanctity foundation has a probability of being selected of about 0.70. In contrast, loyalty and authority tend to be less important for all groups. Here, averaging across all alternatives, respondents have a bit more than a 40% chance of selecting acts associated with loyalty.

The results using the MFV suggest a few interesting conclusions. First, while there are differences across partisan and ideological groups, and while they generally correspond with expectations, they are substantively quite small. Second, the marginal mean for the amoral category is close to zero for all groups, which means that such violations were almost never chosen as more morally wrong, regardless of the moral category associated with the alternative. A key claim in past work is that left-wing groups tend to view the binding foundations as closer to norm violations than moral principles. If this were true, we would expect the marginal means for loyalty, authority, and sanctity, among left-wing groups, to be similar in magnitude to the marginal mean for the amoral acts. But left-wing respondents almost always select acts associated with binding foundations over acts in the amoral category. Indeed, the sanctity foundation has the highest marginal mean for Republicans, Independents, *and* Democrats. Given that acts in the MFV were selected to be roughly comparable in terms of severity, we are limited in what we can say about the relative importance of sanctity outside of this (unrepresentative) set. What we can confidently assert, however, is that all political groups treat sanctity (and other binding) violations as distinct from violations of norms.

Figure 3. Study 2 Estimates for Lucid Survey (MFV versus New Set of Acts)



Notes: Each point represents the mean rank-score for a particular foundation in a particular group. Extended lines are bootstrapped 95% confidence bounds. $\eta^2(\text{founds})$ is the variance in rank-scores explained by the five foundations, and $\eta^2(\text{founds}*\text{group})$ is the additional variance explained by group differences in the pattern of foundation rank-scores.

Turning to the bottom panels of Figure 3, despite a very different research design, we find very similar patterns. First, there is little additional explained variance as a result of adding the interactions of political groups with the moral foundations. The differences in MTurk moral severity ratings for the five foundations alone explain about 27% of the variance in respondents' choices. The additional variance explained by adding the interaction with political groups is close to zero. As with the previous results, some differences in the importance of foundations across groups are observed, and these appear consistent with expectations, but they are small. One possible exception is the difference in the estimated coefficient for sanctity comparing left-wing groups with right-wing groups. For Republicans, a 1-point increase in the difference in the severity of violating alternative A relative to alternative B is associated with a 13-point increase in the probability of selecting alternative A as more morally wrong. For Democrats, by contrast, this increase is about 9 points. The respective values for Conservative and Liberal identifiers are 12 points and 8 points.

Replicating the finding from 2018, however, it is striking to note that sanctity violations are most strongly related to choices for all political groups. Importantly, we allowed for substantially more variation in the severity of violations in this second survey, so these results may be more generalizable than those from the MFV (though more work is required to draw conclusions about this pattern with confidence). Beyond sanctity, care and fairness are the next most influential foundations for choice, with loyalty and authority the least important. Differences in the severity of loyalty violations across acts are insignificantly related to respondents' choices, while differences in severity of authority violations have a small, but significant relationship to choices.

As a whole, Study 2 provides additional evidence that differences in moral priorities across political groups are small. The patterns are similar whether we rely on the MFV or a newly assembled set of acts, and whether we rely on a priori and mutually exclusive coding of the acts within particular foundations, or whether we measure people's perceptions of which foundations the acts violate and to what extent.

General Discussion

Our analyses suggest that differences in moral priorities between the American left and right are smaller than suggested by previous research using the MFQ. These differences in moral priorities are also much smaller than those estimated in past work for political values, such as patriotism, social order, and equality (Jacoby 2014). It is important to reiterate that we do find differences in some instances, and they are consistent with expectations from previous work: the left places relatively less importance on binding foundations than the right (Graham, Haidt, and Nosek 2009; Kivikangas et al. 2021), but these gaps appear too small to sustain the argument that the sharp divisions of contemporary politics are built upon the five foundations of care, fairness, loyalty, authority, and sanctity (Haidt and Graham 2007). In the final pages of the article, we consider possible interpretations of these results and their broader implications.

The primary difference in methodology between our work and previous work using the MFQ is that the latter uses Likert scales, which are independent across moral considerations, while we force respondents to prioritize moral principles. One might argue that our method thus obscures an important point: even if the left and right largely agree on rank-ordering, there are perhaps some moral concerns that the right considers legitimate but the left rejects altogether, and this may be an important source of political conflict. Study 2, however, suggests that the left also treats the binding foundations as legitimate moral concerns, consistently distinguishing them

from the domain of personal preference and convention. Moreover, our 2022 data in Study 2 estimates the size of the relationship between violations of particular foundations and choice, controlling for the other foundations, and finds similar results. But even if we accept this critique, our results still represent a qualification of previous work. Moral judgment often involves conflict among moral principles, and it is critical to understand how citizens make tradeoffs.

Importantly, however, we think there are reasons to be skeptical about the estimates of absolute importance implied by the MFQ. First, independent relevance or Likert scales may be unreliable methods for eliciting the personal importance of considerations for decision making. In the domain of voting behavior, for example, research suggests that ratings of the personal importance of political issues are largely unrelated to the weight that citizens implicitly place on these issues when choosing between hypothetical candidates (Bartle and Laycock 2012, Leeper and Robison 2018). Clifford, Iyengar, et al. (2015, 1179) make this point explicitly with respect to the MFQ: “As Graham et al. argue (2009, p.1031), moral relevance ‘does not necessarily measure how people actually make moral judgments,’ but these ratings are ‘best understood as self-theories about moral judgment.’ Yet, individuals’ theories of morality (i.e., endorsement of moral principles) might diverge from their specific moral judgments (Haidt, 2011).” One reason may be that respondents are reluctant to admit that considerations highlighted by the researcher, or that are salient in some broader context (e.g., the national media, one’s group), are unimportant to them. Another reason may be that respondents fail to consider the possibility of conflict among considerations unless forced by the structure of the survey item.

There is also a question of how to interpret citizens’ judgments regarding the relevance of the foundations in the MFQ. In contrast to MFT, Gray and colleagues argue that moral judgment

is fundamentally about harm. Their theory of dyadic morality suggests people map actions to a fuzzy cognitive template of an intentional agent causing damage to a vulnerable patient—the closer the fit, the more wrong an act is perceived to be. An alternative interpretation of the MFQ is that right-wing citizens tend to believe that actions that disrespect authority, show disloyalty to the group, or degrade the self, map—if only imperfectly—to this more general schema (Schein and Gray 2015). In related work, Stanley, Yin, and Sinnott-Armstrong (2019) find that moral judgments regarding harmless taboo violations are rooted in perceptions of the ex-ante likelihood that the act would cause harm. They argue that people are making moral judgments based on expectations of the consequences of the act before it is committed. That is, acts that can reasonably be expected to cause harm are immoral, even if no one is ultimately harmed. In this way, the MFQ may be tapping citizens' beliefs about the value of various cues for deciding whether an act is likely to have negative consequences, rather than beliefs about the importance of fundamental moral principles. To be clear, our evidence is merely consistent with the latter perspective, and much more research is needed on this topic.

As noted in the introduction, however, left-right conflict may also be “baked-in” to the MFQ. That is, the items measuring the importance of the binding foundations are biased toward right-wing concerns, such as sexual behavior, religion, and patriotism (Schein and Gray 2018). It is possible that a revised MFQ that was politically neutral with respect to content—or which balanced right-wing concerns with left-wing concerns in the binding domain—would produce less disagreement. Related work using both experimentation and panel data provides evidence that responses to the MFQ may be partly endogenous to political identity (Ciuk 2018, Hatemi, Crabtree, and Smith 2019).

Our work is complementary to this line of research in the sense that it provides additional reasons to be skeptical that moral disagreement is the prime mover of contemporary political conflict. An important limitation of past work on the causal ordering of moral foundations and ideology (Ciuk 2018, Hatemi, Crabtree, and Smith 2019) is that it relies on the MFQ which, as we have discussed, has significant drawbacks. Another limitation is that, by assumption, research using panel data or randomly assigned group cues can only investigate changes in moral priorities and ideology at the margin. It thus leaves open the possibility that partisan and ideological gaps in *preexisting* moral priorities are due, in large part, to the causal influence of morality on politics. This could be true if the two are largely in equilibrium; for example, if moral priorities have a large influence on political orientations during “critical periods” of early adulthood, and the subsequent relationship is just short-term variation around these equilibrium values. Moreover, even if moral priorities are largely rooted in broad political orientations, they may still have a causal influence on other phenomena of interest, such as the nature of political discourse, degree of animosity toward political out-groups, or decision making in contexts without explicit partisan or ideological cues. Importantly, then, we provide evidence that—*regardless of origin*—the moral priorities posited by MFT may not sharply divide citizens of the left and right after all. As noted in the introduction, this implies greater optimism about the possibility for reasoned discourse across partisan and ideological lines.

Yet, if we are right about the extent of moral *agreement* in American politics, we are left with a puzzle. Contemporary American politics exhibits several features, such as intense feelings and unwillingness to compromise, that suggest political attitudes are rooted in core moral principles (i.e., “moral conviction”; Ryan 2017, Skitka and Morgan 2014). How can we reconcile moral agreement with what appears to be intense moral conflict? One answer is that moral

conviction and moral disagreement are separable constructs: it is possible for the left and right to largely agree on the relative importance of the five foundations (on average) but disagree on the implications of these for political issues. First, citizens may disagree about the *applicability* of a principle to an issue. People who are “pro-life” on abortion may believe it is an issue that revolves around considerations of care and harm, while “pro-choice” folks may see it as issue of gender equality and fairness. In such cases, it may be possible to reduce animosity through discourse aimed at elaborating the reasons for opponents’ positions and thus appreciation for, if not agreement with, those positions.

Second, even if citizens agree on both the importance of moral principles and their relevance to a given issue, they may nonetheless apply the same moral principle in different ways. For example, people who wish to restrict immigration may believe that the harm to Americans would outweigh the benefits to immigrants, while pro-immigrant individuals may believe the opposite. And as with the work of Schein and Gray (2015) and Stanley, Yin, and Sinnott-Armstrong (2019), there may be disagreement over what kinds of actions cause harm. Regardless of who is right or wrong, this would constitute a difference in beliefs about relevant facts, not a fundamental conflict over the importance of moral principles.

Finally, the literature suggests other possible explanations for the intensity and vitriol of contemporary American politics that do not require sharp moral divisions between the left and right, such as racial and ethnic change (Jardina 2019) and economic change (Autor et al. 2016).

To conclude, we emphasize that our paper is not a test of MFT itself, which is a general framework for understanding human morality, and does not rise nor fall based on our findings here. Indeed, individual differences in the importance of the proposed foundations are only one aspect of this framework and applications to American politics are not themselves tests of the

underlying theory.⁷ While other recent work confronts MFT itself (Curry 2016, Schein and Gray 2018), our interest in the theory arises because it is the primary vehicle for documenting the extent of moral conflict in the United States. Thus, an important limitation of our study is that we cannot speak to the possibility that an alternative framework for investigating human morality would reveal larger gaps between the left and right. For example, Curry, Chesters and Van Lissa (2019) have recently proposed an alternative set of basic moral values along with an instrument for measuring these values. As another example, Emler (2003) argues that different “styles” of moral reasoning (e.g., social order vs principled reasoning styles) are used primarily as a vehicle for expressing political identities. Given the influence of MFT in the study of contemporary American political conflict, however, our results remain an important contribution.

References

- Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi. 2016. "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure." NBER Working Paper 22637, National Bureau of Economic Research.
doi:<http://www.nber.org/papers/w22637>.
- Bartle, John, and Samantha Laycock. 2012. "Telling More Than They Can Know? Does the Most Important Issue Really Reveal What Is Most Important to Voters?" *Electoral Studies* 31 (4):679-688.

⁷ However, our covariance structure analyses, along with other recent work, do suggest that MFT's primary instrument, the MFQ, does not provide support for one of the theory's key claims, namely, a five (or more) factor structure for morality.

- Ciuk, David J. 2018. "Assessing the Contextual Stability of Moral Foundations: Evidence from a Survey Experiment." *Research & Politics* 5 (2):2053168018781748.
- Ciuk, David J. and William G. Jacoby. 2015. "Checking for Systematic Value Preferences Using the Method of Triads." *Political Psychology* 36 (6): 709-728.
- Clifford, Scott, Vijeth Iyengar, Roberto Cabeza, and Walter Sinnott-Armstrong. 2015. "Moral Foundations Vignettes: A Standardized Stimulus Database of Scenarios Based on Moral Foundations Theory." *Behavior research methods* 47 (4):1178-1198.
- Clifford, Scott, and Jennifer Jerit. 2013. "How Words Do the Work of Politics: Moral Foundations Theory and the Debate over Stem Cell Research." *The Journal of Politics* 75 (3):659-671.
- Coppock, Alexander, and Oliver A McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research & Politics* 6 (1):1-14.
- Curry, Oliver Scott. 2016. "Morality as Cooperation: A Problem-Centred Approach." In *The Evolution of Morality*, edited by Todd K. Shackelford and Randal D. Hansen, 27-51. Cham: Springer International Publishing.
- Curry, Oliver Scott, Matthew Jones Chesters and Casper J. Van Lissa. 2019. "Mapping Morality with a Compass: Testing the Theory of 'Morality-as-cooperation' with a New Questionnaire." *Journal of Research in Personality* 78: 106-124.
- Ditto, Peter H, and Spassena P Koleva. 2011. "Moral Empathy Gaps and the American Culture War." *Emotion Review* 3 (3):331-332.
- Emler, Nicholas. 2003. "Morality and Political Orientations: An Analysis of Their Relationship." *European Review of Social Psychology* 13 (1): 259-291.

- Federico, Christopher M, Christopher R Weber, Damla Ergun, and Corrie Hunt. 2013. "Mapping the Connections between Politics and Morality: The Multiple Sociopolitical Orientations Involved in Moral Intuition." *Political Psychology* 34 (4):589-610.
- Feinberg, Matthew, and Robb Willer. 2013. "The Moral Roots of Environmental Attitudes." *Psychological Science* 24 (1):56-62.
- Feinberg, Matthew and Rob Willer. 2015. "From Gulf to Bridge: When Do Moral Arguments Facilitate Political Influence?" *Personality and Social Psychology Bulletin* 41 (12):1665-1681.
- Frimer, Jeremy A. 2020. "Do Liberals and Conservatives Use Different Moral Languages? Two Replications and Six Extensions of Graham, Haidt, and Nosek's (2009) Moral Text Analysis." *Journal of Research In Personality* (84): 103906.
- Graham, Jesse, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P Wojcik, and Peter H Ditto. 2013. "Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism." In *Advances in Experimental Social Psychology*, 55-130. Elsevier.
- Graham, Jesse, Jonathan Haidt, and Brian A Nosek. 2009. "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology* 96 (5):1029-1046.
- Graham, Jesse, Brian A Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter H Ditto. 2011. "Mapping the Moral Domain." *Journal of Personality and Social Psychology* 101 (2):366.
- Haidt, Jonathan. 2008. "Morality." *Perspectives on Psychological Science* 3 (1):65-72.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York, NY: Vintage.

- Haidt, Jonathan, and Jesse Graham. 2007. "When Morality Opposes Justice: Conservatives Have Moral Intuitions That Liberals May Not Recognize." *Social Justice Research* 20 (1):98-116.
- Hainmueller, Jens, Daniel J Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices Via Stated Preference Experiments." *Political Analysis* 22 (1):1-30.
- Hatemi, Peter K., Charles Crabtree, and Kevin B. Smith. 2019. "Ideology Justifies Morality: Political Beliefs Predict Moral Foundations." *American Journal of Political Science* 63 (4):788-806.
- Iurino, Kathryn, and Gerard Saucier. 2018. "Testing Measurement Invariance of the Moral Foundations Questionnaire across 27 Countries." *Assessment*:1073191118817916.
- Jacoby, William G. 2014. "Is There a Culture War? Conflicting Value Structures in American Public Opinion." *American Political Science Review* 108 (4):754-771.
- Jardina, Ashley. 2019. *White Identity Politics*: Cambridge University Press.
- Johnson, Kate M, Ravi Iyer, Sean P Wojcik, Stephen Vaisey, Andrew Miles, Veronica Chu, and Jesse Graham. 2014. "Ideology-Specific Patterns of Moral Indifference Predict Intentions Not to Vote." *Analyses of Social Issues and Public Policy* 14 (1):61-77.
- Kertzer, Joshua D, Kathleen E Powers, Brian C Rathbun, and Ravi Iyer. 2014. "Moral Support: How Moral Values Shape Foreign Policy Attitudes." *The Journal of Politics* 76 (3):825-840.
- Kivikangas, J Matias, Belén Fernández-Castilla, Simo Järvelä, Niklas Ravaja, and Jan-Erik Lönnqvist. 2021. "Moral Foundations and Political Orientation: Systematic Review and Meta-Analysis." *Psychological Bulletin* 147 (1): 55.

- Koleva, Spassena P, Jesse Graham, Ravi Iyer, Peter H Ditto, and Jonathan Haidt. 2012. "Tracing the Threads: How Five Moral Concerns (Especially Purity) Help Explain Culture War Attitudes." *Journal of Research in Personality* 46 (2):184-194.
- Leeper, Thomas J, and Joshua Robison. 2018. "More Important, but for What Exactly? The Insignificant Role of Subjective Issue Importance in Vote Decisions." *Political Behavior*:1-21.
- Monroe, Andrew E., James B. Wyngaarden, III and E. Ashby Plant. 2021. "'They Should Have Followed the Rules': Trade-offs Between Fairness and Authority Values Predict Judgments of Social Justice Protests." *Social Psychological and Personality Science* 12 (4): 471-478.
- Neiman, Jayme L., Frank J. Gonzalez, Kevin Wilkinson, Kevin B. Smith and John R. Hibbing. 2016. "Speaking Different Languages or Reading from the Same Script? Word Usage of Democratic and Republican Politicians." *Political Communication* 33 (2): 212-240.
- Nilsson, Artur, and Arvid Erlandsson. 2015. "The Moral Foundations Taxonomy: Structural Validity and Relation to Political Ideology in Sweden." *Personality and Individual Differences* 76:28-32.
- Ryan, Timothy J. 2017. "No Compromise: Political Consequences of Moralized Attitudes." *American Journal of Political Science* 61 (2):409-423.
- Schein, Chelsea, and Kurt Gray. 2015. "The Unifying Moral Dyad: Liberals and Conservatives Share the Same Harm-Based Moral Template." *Personality and Social Psychology Bulletin* 41 (8):1147-1163.

- Schein, Chelsea, and Kurt Gray. 2018. "The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm." *Personality and Social Psychology Review* 22 (1):32-70.
- Schwartz, Shalom H. 1992. "Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries." In *Advances in Experimental Social Psychology*, 1-65. Elsevier.
- Skitka, Linda J, and G Scott Morgan. 2014. "The Social and Political Implications of Moral Conviction." *Political Psychology* 35:95-110.
- Smith, Colin Tucker, Kate A Ratliff, Liz Redford, and Jesse Graham. 2019. "Political Ideology Predicts Attitudes toward Moral Transgressors." *Journal of Research in Personality* 80:23-29.
- Smith, Kevin B., John R. Alford, John R. Hibbing, Nicholas G. Martin, and Peter K. Hatemi. 2017. "Intuitive Ethics and Political Orientations: Testing Moral Foundations as a Theory of Political Ideology." *American Journal of Political Science* 61 (2):424-437.
- Stanley, Matthew L., Siyuan Yin, and Walter Sinnott-Armstrong. 2019. "A Reason-Based Explanation for Moral Dumbfounding." *Judgment and Decision Making* 14 (2):120-129.
- Sterling, Joanna, and John T. Jost. 2018. "Moral Discourse in the Twitterverse: Effects of Ideology and Political Sophistication on Language Use Among U.S. Citizens and Members of Congress." *Journal of Language and Politics* 17 (2): 195-221.
- Watkins, Leah, Robert Aitken, and Damien Mather. 2016. "Conscientious Consumers: A Relationship between Moral Foundations, Political Orientation and Sustainable Consumption." *Journal of Cleaner Production* 134:137-146.

Waytz, Adam W., James Dungan and Liane Young. 2013. "The Whistleblower's Dilemma and the Fairness–Loyalty Tradeoff." *Journal of Experimental Social Psychology* 49 (6):1027-1033.

Weinschenk, Aaron C, and Christopher T Dawes. 2019. "Moral Foundations, System Justification, and Support for Trump in the 2016 Presidential Election." *The Forum*.

Winterich, Karen Page, Yinlong Zhang, and Vikas Mittal. 2012. "How Political Identity and Charity Positioning Increase Donations: Insights from Moral Foundations Theory." *International Journal of Research in Marketing* 29 (4):346-354.

SUPPLEMENTAL APPENDIX

Contents

- A. [The Moral Foundations Questionnaire](#)
- B. [Measurement Properties of the MFQ](#)
- C. [Information about Surveys in Studies 1 and 2, Including Definitions](#)
- D. [Moral Vignettes for 2018 Lucid Survey](#)
- E. [Moral Vignettes for 2022 Lucid Survey](#)
- F. [Alternative Measures of Ideology in Studies 1 and 2](#)
- G. [Supplemental Results](#)
- H. [References](#)

Supplemental Appendix A. The Moral Foundations Questionnaire

The 30-item moral foundations questionnaire was downloaded from the following website on 1/27/2020: <https://moralfoundations.org/questionnaires/>. The labels attached to each item (in bolded red) are our own and used to reference the items in the paper.

Part 1. When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? Please rate each statement using this scale: [0] = not at all relevant (This consideration has nothing to do with my judgments of right and wrong), [1] = not very relevant, [2] = slightly relevant, [3] = somewhat relevant, [4] = very relevant, [5] = extremely relevant (This is one of the most important factors when I judge right and wrong)

- _____ Whether or not someone suffered emotionally (**emote**)
- _____ Whether or not some people were treated differently than others (**treatd**)
- _____ Whether or not someone's action showed love for his or her country (**lovec**)
- _____ Whether or not someone showed a lack of respect for authority (**auth**)
- _____ Whether or not someone violated standards of purity and decency (**purity**)
- _____ Whether or not someone was good at math
- _____ Whether or not someone cared for someone weak or vulnerable (**weak**)
- _____ Whether or not someone acted unfairly (**unfair**)
- _____ Whether or not someone did something to betray his or her group (**betray**)
- _____ Whether or not someone conformed to the traditions of society (**conform**)
- _____ Whether or not someone did something disgusting (**disgust**)
- _____ Whether or not someone was cruel (**cruel**)
- _____ Whether or not someone was denied his or her rights (**rights**)
- _____ Whether or not someone showed a lack of loyalty (**loyal**)
- _____ Whether or not an action caused chaos or disorder (**chaos**)
- _____ Whether or not someone acted in a way that God would approve of (**goddis**)

Part 2. Please read the following sentences and indicate your agreement or disagreement:

[0] [1] [2] [3] [4] [5]

Strongly
disagree

Moderately
disagree

Slightly
disagree

Slightly
agree

Moderately
agree

Strongly
agree

- _____ Compassion for those who are suffering is the most crucial virtue. (**compat**)
- _____ When the government makes laws, the number one principle should be ensuring that everyone is treated fairly. (**treatf**)
- _____ I am proud of my country's history. (**proudc**)
- _____ Respect for authority is something all children need to learn. (**authc**)
- _____ People should not do things that are disgusting, even if no one is harmed. (**nodisgust**)
- _____ It is better to do good than to do bad.
- _____ One of the worst things a person could do is hurt a defenseless animal. (**hurt**)
- _____ Justice is the most important requirement for a society. (**justice**)
- _____ People should be loyal to their family members, even when they have done something wrong. (**loyalf**)
- _____ Men and women each have different roles to play in society. (**roles**)
- _____ I would call some acts wrong on the grounds that they are unnatural. (**notnat**)
- _____ It can never be right to kill a human being. (**kill**)
- _____ I think it's morally wrong that rich children inherit a lot of money while poor children inherit nothing. (**rich**)
- _____ It is more important to be a team player than to express oneself. (**team**)
- _____ If I were a soldier and disagreed with my commanding officer's orders, I would obey anyway because that is my duty. (**obey**)
- _____ Chastity is an important and valuable virtue. (**chaste**)

Supplemental Appendix B. Measurement Properties of the MFQ

The MFQ (see Supplemental Appendix A) asks respondents to rate the relevance of different moral concerns to their judgments of right and wrong (Form A) and also about their agreement with statements that reflect the importance of such concerns (Form B). While the MFQ has inarguably made a valuable contribution to the literature on mass politics, it has poor measurement properties. To demonstrate this, we consider three national surveys with adult respondents in the United States: a collaborative module on the 2012 Cooperative Congressional Election Study, the 10th wave of The American Panel Study, and the fourth wave of the National Survey of Youth and Religion (see Table B1). The first contains the full 30-item MFQ while the latter two contain the 20-item short-version.

We estimate (via full-information maximum likelihood) two covariance structure models for each dataset. In the first, the MFQ items are modeled as a linear function of their respective (latent) moral foundations, the correlations among the five latent foundations are freely estimated, and the error terms for the items are assumed to be independent and identically distributed. This is the preferred specification of Graham et al. (2011) in their MFQ-validation study. The second model is identical, except we add two additional latent variables, one for each form of the MFQ (A and B). That is, we model the items as a function of both their respective latent moral foundation *and* a latent variable that represents individual differences in respondents' use of the measurement instrument for that item. We refer to these as “methods” factors because they represent variance in responses that is independent of content and due solely to the way the questions are asked.

Figure B1A shows estimates for the first set of models (without methods factors) and Figure B1.B shows the second set (with methods factors). Plotted symbols represent standardized factor loadings of each item on its respective moral foundation. The small black points in Figure B1.b represent the standardized loading of each item on its respective methods factor. We turn first to the models without methods factors. There are a few noteworthy features of these results. First, fit for the canonical MFT model is low for all three surveys using both the root mean squared error of approximation (0.07, 0.12, 0.10, for CCES, TAPS, and NSYR, respectively) and the Tucker-Lewis Index (0.67, 0.59, 0.63). The former is a measure of absolute fit, while the latter is a measure of comparative fit to a baseline model which assumes uncorrelated observed indicators. The TLI thus better captures situations with many poor indicators of the latent variables of interest. Looking at the top three panels of Figure B1A, this appears to be the case in our data. Across the datasets, many of the item loadings fall below 0.40. Further, with the exception of the sanctity foundation, there is a clear pattern to the loadings: items using Form B—the agree/disagree format (the second three items listed in each foundation)—have consistently smaller loadings than those using Form A (the first three items in each foundation). This suggests that measurement variance may be quite important to the MFQ.

This is confirmed in models that add methods factors. First, model fit improves substantially (RMSEA = (0.05, 0.08, 0.05) and TLI = (0.82, 0.81, 0.91)). Second, as shown in Figure B1B, many of the items that appear to be good indicators in the model without methods factors are poor indicators once we account for methods variance. The standardized loadings of the items on their respective methods factors reinforce these findings. In every case they are non-negligible, and in many cases—especially for Form A—variance in items attributable to instrumentation exceeds that attributable to moral content. That is, many items are better

measures of individual differences in respondents' use of the survey instrument than individual differences in moral priorities.

These results also suggest that the MFQ provides only weak support for the five-factor model of morality. We report the estimated correlations of the latent moral foundations for each survey in Figure B2. Looking first at models without methods factors, the correlations among the latent factors are strikingly large and this is especially true for foundations that share a superordinate domain (i.e., individualizing or binding), for which the correlations often approach unity. For example, in the CCES data, care and fairness are correlated at 0.96 and authority and loyalty are correlated at 0.98. Indeed, the covariance matrix of the latent factors for models using the CCES and TAPS are not positive-definite, which is consistent with other recent work (Iurino and Saucier 2018).

The correlations among foundations that do *not* share a superordinate domain change drastically once we account for measurement variance. In the CCES and TAPS, these interdomain correlations fall toward zero in each case, while the intradomain correlations remain very high and, in some cases, increase in magnitude. In the NSYR, the interdomain correlations remain high but *switch sign*. That is, individualizing and binding foundations are often strongly and negatively correlated with one another. For example, loyalty and fairness are correlated at -0.65. Overall, these results strongly suggest that a five-factor model is more complex than necessary. At a bare minimum, the MFQ is not strongly differentiating moral content *within* the individualizing and binding domains, and the NSYR even suggests substantial relationships across superordinate domains.

Recent work finds a similar pattern of large correlations within the superordinate dimensions and only a modest increase in fit from a two to a five-factor model (Davies, Sibley, and Liu 2014, Graham et al. 2011, Iurino and Saucier 2018, Nilsson and Erlandsson 2015, Yilmaz et al. 2016). We add to this literature by drawing on three national U.S. surveys that have not previously been examined (indeed, most of this research has used data in other contexts). We also extend this work by demonstrating that the MFQ items are strongly contaminated by measurement variance.

Table B1. Information about MFQ Datasets

2012 CCES

The 2012 Cooperative Congressional Election Study (CCES) dataset was collected as part of a collaborative module on the 2012 CCES (Ansolabehere and Schaffner 2013a). The 2012 CCES consisted of 48 teams each of which purchased 1,000 respondents. Half of the survey for each team contained team-specific content and half was common to all teams (the “common content”). The pre-election wave of the study was conducted during October and the post-election wave was conducted over the two weeks following the election. All surveys were conducted over the internet by YouGov.

YouGov attempts to mimic a random sample of the target population through a three-stage process (Ansolabehere and Schaffner 2013b). First, a stratified (by age, race, gender, and education) random sample was drawn from the sampling frame for the target population. The frame was constructed from U.S. citizens in the 2008 American Community Survey matched to voter registration and turnout in the November 2008 Current Population Survey and religion, church attendance, born again or evangelical status, news interest, party identification and ideology from the 2008 Pew U.S. Religious Landscape Survey. Second, the closest matching respondent in the pool of opt-in panel respondents (in this case, the pool is YouGov’s PollingPoint Panel along with the E-Rewards and Western Wats panels) was found for each member of the target population sample. The matching function used gender, age, race, Census region, education, news interest, marital status, partisanship, ideology, religious identification, religiosity, income, voter registration, and metropolitan residence status. Finally, the matched cases were weighted to the constructed sampling frame using propensity scores based on age, education, gender, and turnout. Sample characteristics are shown in Figure B3.

2012 TAPS

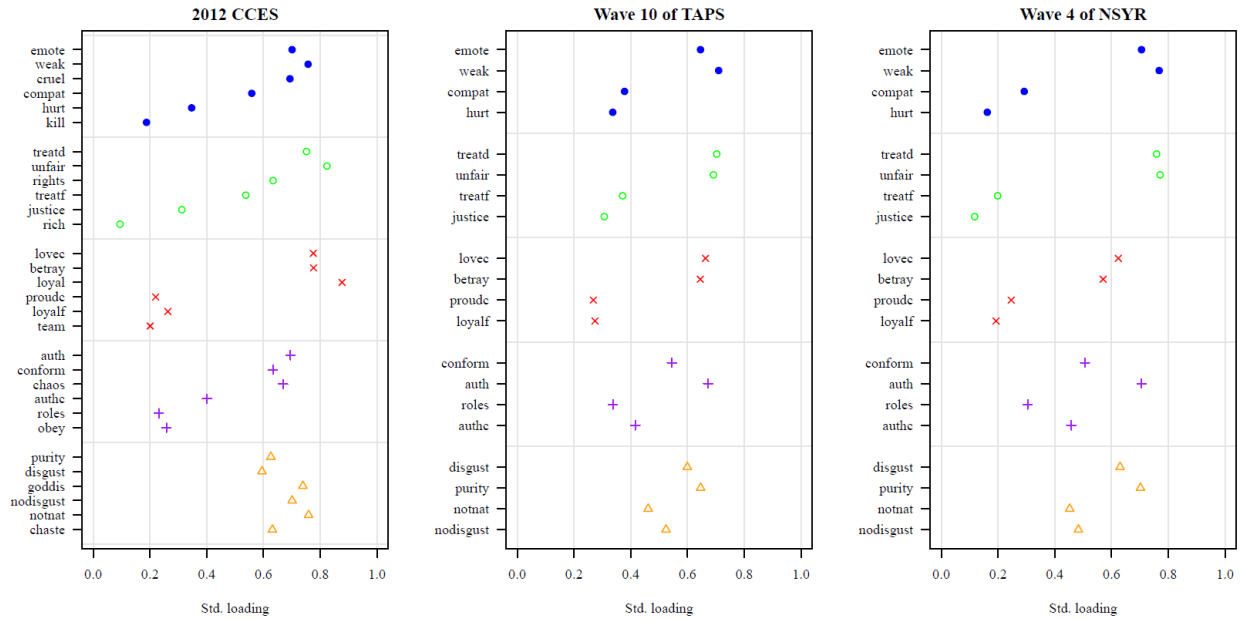
The American Panel Survey (TAPS) is a monthly panel study that began surveying respondents in Fall 2011. The sample is intended to be representative of U.S. adults and was constructed by stratified (by age and ethnic group) random sampling of residential addresses in all 50 states. A panel of more than 2,000 adults was initially recruited. TAPS surveys were administered online by GfK/Knowledge Networks. Panelists who did not have a computer or online service were provided a computer and internet access by TAPS. We rely on the tenth wave of TAPS conducted in September of 2012. For additional details on TAPS, see <https://wc.wustl.edu/american-panel-survey>. Sample characteristics are shown in Figure B4.

Fourth Wave of the NSYR

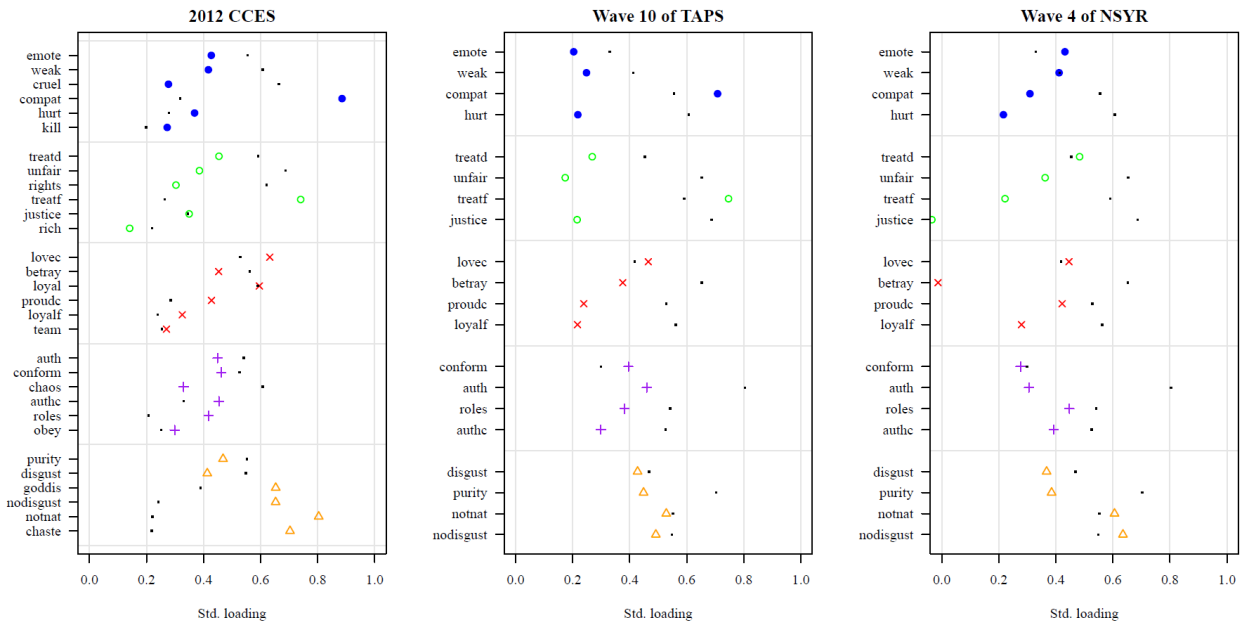
The National Study of Youth and Religion is a four-wave panel study. The first wave was conducted in 2002 and is based on a random-digit-dialing sample of households (with in-house subject randomization) with youth between the ages of 13 and 17 (3,370 completed cases). The second wave (in 2005) successfully reinterviewed 78% of the original sample; the third wave (in 2007-2008) reinterviewed 77% of the original sample; and the fourth wave (in 2012) reinterviewed 67% of the original sample (85% of these took the survey online and 15% over the phone). The moral foundations questionnaire was included in the fourth wave, when respondents were aged 23-28. More information about the NSYR can be found at the following website: <https://youthandreligion.nd.edu/>. Sample characteristics are shown in Figure B5.

Figure B1. Standardized Factor Loadings of MFQ Items

A. Models without Methods Factors



B. Models with Methods Factors



Notes: The first (second) set of three or two items in each foundation are from Form A (B). The small black dots are standardized loadings of each item on its respective methods factor. The order of moral foundations (from top to bottom) is care, fairness, loyalty, authority, sanctity.

Figure B2. Correlations Among the Five Latent Foundations in the MFQ



Notes: Entries in the lower (upper) triangle of each table are estimates from models without (with) methods factors. Bolded entries equal or exceed 0.40 and italicized entries are below 0.40.

Figure B3. Sample Characteristics for CCES

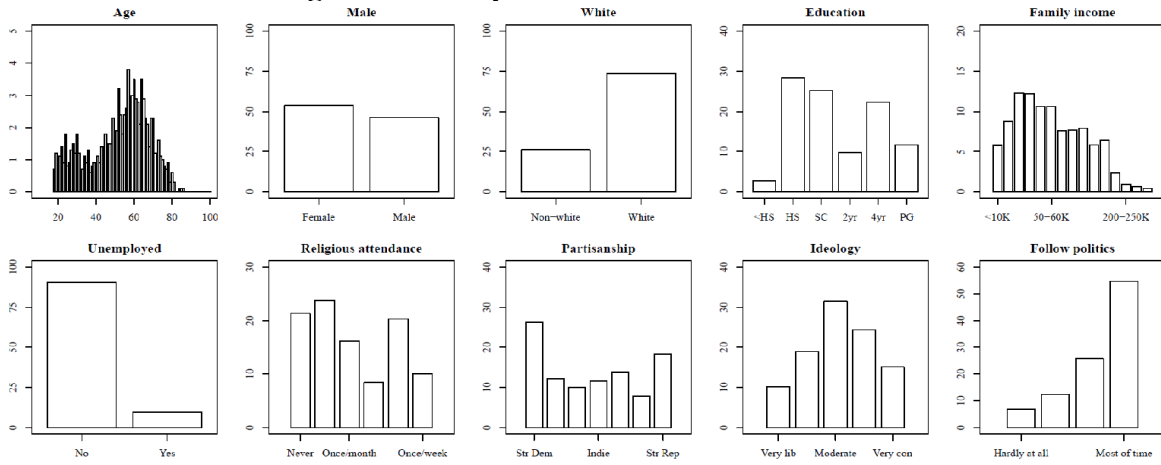


Figure B4. Sample Characteristics for TAPS

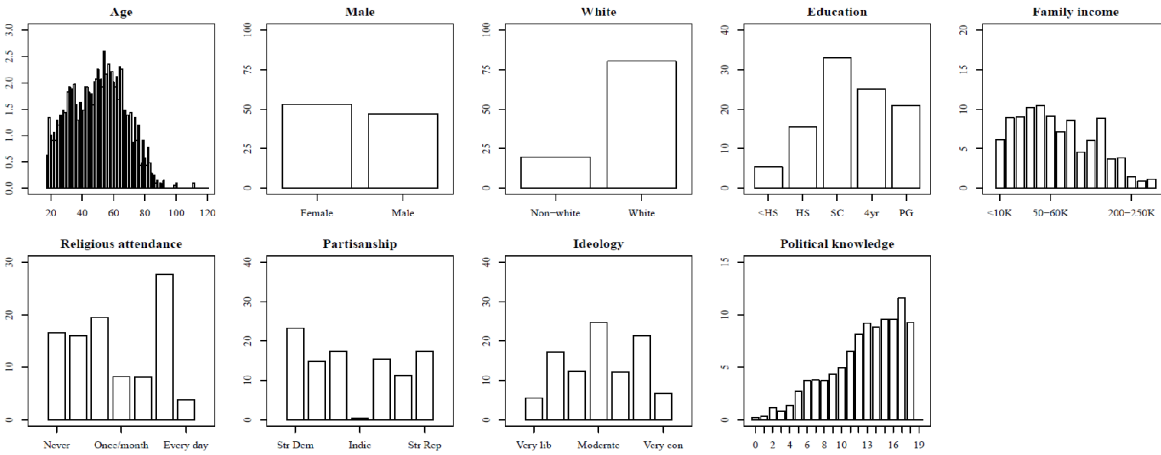
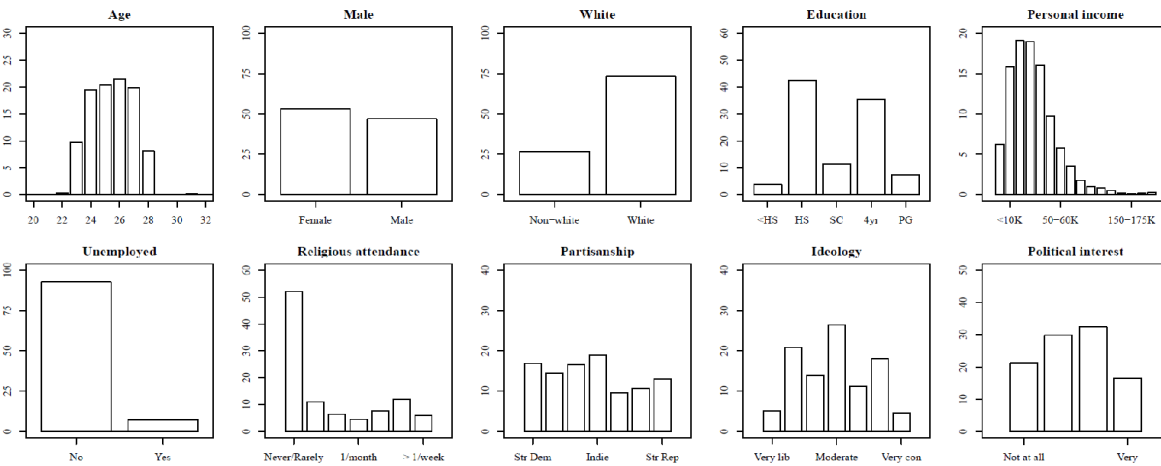


Figure B5. Sample Characteristics for NSYR



Supplemental Appendix C. Information about Surveys in Studies 1 and 2

2017 Mechanical Turk Survey

Data collection for the Mechanical Turk sample took place in April 2017. In order to qualify to participate in the study, MTurk workers needed to be residents of the United States, have a 95% approval rating, and have completed 500 HITs. In total, 406 respondents completed the study, and each was paid \$0.50. Descriptive statistics, including demographics and data regarding political affiliation, can be found in Table C1.

2019 Dynata Survey

We recruited a second sample through Dynata (formerly Survey Sampling International), which is a market research firm that specializes in internet surveys. Dynata recruits potential respondents through various online communities, social networks, and website ads. They also make efforts to reach groups such as racial minorities and seniors. Potential participants are screened, then invited to the panel. For a particular survey, Dynata randomly selects from its pool of respondents. The present study did not use quotas, but the authors requested a sample to match the US Census (age 18+), as closely as possible, on age, gender, race, income, and education. Demographics and data regarding political affiliations for this sample is shown in Table C1.

2018 Lucid Survey

We collected a non-probability sample using Lucid Marketplace which resulted in 990 completed surveys. Lucid is a platform which links researchers with over 250 sample providers. Providers direct their panelists to Lucid’s Marketplace and Lucid directs these panelists to available surveys for which they qualify. In the United States, there were approximately 17.5 million unique visitors to Marketplace during the sixth-month period preceding January 2019 (personal communication with Lucid, 2019). Lucid determines unique respondents within each survey through a combination of IP address, a unique Lucid identification number, and a unique panel identification number (Lucid, private correspondence). In a recent study, Coppock and McClellan (2019) find (with one exception) that samples drawn from Lucid replicate previously published experimental findings and show effect sizes comparable to both the original studies and samples drawn from MTurk.

The survey was fielded between August 14th and August 17th, 2018. We employed quotas throughout our recruitment process to obtain distributions of age, gender, race, ethnicity, and region that approximate that of the general U.S adult population (based on adults in the 2016 American Community Survey). For example, we aimed to have 75% of respondents in our sample identify as White, around 11% between the ages of 18 to 24, and 49% identifying as Male. While we were able to meet quotas on race and gender, our sample does tend to be older. Prior to the moral decision task, we employed two attention-check screeners that terminated those respondents who failed to answer either item correctly. We employed these attention check screeners to account for low-effort respondents who are likely to provide low-quality data that can bias results (Alvarez et al. 2019). Specifically, we provided respondents with a set of images, and asked them to identify each image that contained a stop-sign (this item also served as a “bot-check”). The other attention-check item simply required respondents to correctly identify the current President of the United States. Respondents who failed either of these items were immediately terminated from the survey and did not participate in the forced-choice task. More

information on our analytical sample’s demographic characteristics can be found below in Table C2.

2022 Lucid Survey

We collected a non-probability sample using Lucid Marketplace which resulted in 916 completed surveys (see the previous discussion of the 2018 Lucid Survey for more information about Lucid). The survey was fielded in March and April of 2022. We employed quotas throughout our recruitment process to obtain distributions of age, gender, race, ethnicity, and region that approximate that of the general U.S adult population (based on adults in the 2016 American Community Survey). More information on our analytical sample’s demographic characteristics can be found below in Table C2.

We employed three attention-check screeners that terminated those respondents who failed to answer correctly. First, we provided respondents with a set of images, and asked them to identify each image that contained a stop-sign (this item also served as a “bot-check”). The other attention-check items required respondents to (1) select a particular response on a survey item and (2) check two specific boxes on a multiple response item. Respondents who failed any of these checks were immediately terminated from the survey and do not appear in our analytical sample. We additionally removed observations from the analytical sample if they had duplicate Lucid respondent IDs in the data, had Qualtrics reCAPTCHA scores less than 0.5, Qualtrics relevant ID duplicate scores of greater than or equal to 75, or Qualtrics fraud scores greater or equal to 30 (the threshold values are those recommended by Qualtrics). Table C2 provides the characteristics of the final analytical sample.

Definitions and Instructions for all Surveys

In each survey, respondents completed an importance ranking task using the method of triads. They were first shown definitions of the five foundations. They were then shown a series of 10 screens, each of which displayed a set of three of the five foundations, and were asked to indicate their most and least important foundation in each set. The moral foundation definitions were identical for the MTurk and Dynata samples (see Table C3). In the Lucid sample, half of respondents (N=888) were randomly assigned to receive the same definitions as the MTurk and Dynata samples, and half (N=889) were randomly assigned to receive a new set of definitions.

Table C1. Characteristics of the 2017 MTurk and 2019 Dynata Surveys

	DYNATA (n = 1478)	MTurk (n = 406)
GENDER (%)		
Male	54	54
Female	45	46
Non-binary/NA	1	
RACIAL GROUP (%)		
White, non-Hispanic	65	86
Black, non-Hispanic	16	8
Hispanic	11	6
Other	8	
HOUSEHOLD INCOME (%)		
Less than \$25k	20	17
\$25k – \$49,999	25	31
\$50k – \$74,999	17	23
\$75k – \$99,999	14	17
\$100k - \$124,999	8	6
\$125k - \$149,999	6	3
\$150k - \$174,999	4	2
\$175k or more	6	1
EDUCATION LEVEL (%)		
Less than HS Diploma	2	< 1
HS Diploma	20	15
Some College	22	23
2-year Degree	11	14
4-year Degree	25	35
Some Postgrad Training	4	11
Postgraduate Degree	15	1
RELIGION (%)		
Protestant	21	21
Catholic	24	14
Other Christian	17	13
Muslim	2	< 1
Jewish	3	3
Other	8	4
No religious affiliation	24	44
AGE GROUP (%)		
18-24	11	NA
35-34	18	NA
35-44	17	NA

Supplemental Appendix to “Competing Moral Minds”, p.6

45-54	17	NA
55-64	19	NA
65-74	18	NA
75+	< 1	NA
REGION (%)		
South	35	NA
Northeast	24	NA
Midwest	20	NA
West	21	NA
POLITICAL PARTY (%)		
Democrat	44	52
Republican	35	34
Independent/Other	21	14
IDEOLOGY (%)		
Liberal	38	50
Conservative	34	30
Moderate	28	20

Table C2. Characteristics of the 2018 and 2022 Lucid Surveys

		2018	2022
Gender			
	Male	51	41
	Female	49	59
Party ID			
	Rep	39	39
	Ind	17	16
	Dem	44	45
Ideo ID			
	Con	34	36
	Mod	35	33
	Lib	31	31
Race			
	Non-Hispanic White	75	79
Political attention			
	A great deal	26	24
	A lot	19	22
	Moderate	26	28
	A little	21	19
	None	7	6
Income			
	< 30K	36	27
	30K - 60K	32	29
	60K - 100 K	21	22
	> 100K	11	21
Education			
	HS or less	30	22
	Some college/2-year	39	35
	4-year	21	27
	Post-grad	10	15
Age			
	18-24	8	4
	24-44	27	29
	45-64	39	35
	65+	25	31

Table C3: Definitions for Foundations

“OLD” DEFINITIONS (MTurk, Dynata, and Lucid sample 1)

We'd like to ask you about some things that are often considered important in determining whether someone's actions are right or wrong. Below are the five concepts that some people think are most important:

HARM/CARE, which involves being sensitive to the suffering of others. For example, when you're deciding whether an action is right or wrong, do you consider whether someone cared for the weak and vulnerable, or whether they showed compassion for those who are suffering?

FAIRNESS/RECIPROCITY, which involves making sure everyone is treated equally. For example, when you're deciding whether an action is right or wrong, do you consider whether someone was treated unfairly, or whether someone acted unfairly?

INGROUP/LOYALTY, which involves allegiance to a group, like your family, your close friends, or your country. For example, when you're deciding whether an action is right or wrong, do you consider whether someone showed love for his or her country, or whether they did something to betray their group?

AUTHORITY/RESPECT, which involves having a good deal of respect for authority. For example, when you're deciding whether an action is right or wrong, do you consider whether someone showed a lack of respect for authority, or whether they teach their kids to be respectful to adults and other authority figures?

PURITY/SANCTITY, which involves upholding standards of purity and decency. For example, when you're deciding whether an action is right or wrong, do you consider whether someone acted in a disgusting way (even if no one was harmed), or whether the act was wrong on the grounds that it was unnatural?

“NEW” DEFINITIONS (Lucid sample 2)

We'd like to ask you about some things that are often considered important in determining whether someone's actions are right or wrong. Below are the five concepts that some people think are most important:

CARE/HARM, which involves being sensitive to the suffering of others. For example, when you're deciding whether an action is right or wrong, do you consider whether someone caused others to feel physical or emotional pain, or whether they showed compassion for those who are suffering?

FAIRNESS/CHEATING, which involves being honest and just. For example, when you're deciding whether an action is right or wrong, do you consider whether someone took advantage of others, or whether everyone was treated equally?

LOYALTY/BETRAYAL, which involves allegiance to those close to you. For example, when you're deciding whether an action is right or wrong, do you consider whether someone betrayed others, or whether they showed devotion to their group?

AUTHORITY/SUBVERSION, which involves respecting status differences. For example, when you're

deciding whether an action is right or wrong, do you consider whether someone was disobedient, or whether they were respectful to their superiors?

SANCTITY/DEGRADATION, which involves upholding standards of purity and decency. For example, when you're deciding whether an action is right or wrong, do you consider whether someone behaved in a disgusting way, or whether they showed respect for the sacred?

Appendix D. Full List of Moral Foundation Vignettes for 2018 Lucid Survey (Study 2)

<p><u>Care-Physical</u></p>	<p>"You see a woman swerving her car in order to intentionally run over a squirrel.", "You see a woman throwing her cat across the room for scratching the furniture.", "You see a teacher hitting a student's hand with a ruler for falling asleep in class.", "You see a woman throwing a stapler at her colleague who is snoring during her talk.", "You see a woman slapping another woman who she is arguing with in the parking lot."</p>
<p><u>Care-Emotion</u></p>	<p>"You see a teenage boy chuckling at an amputee he passes by while on the subway." "You see a girl laughing at another student forgetting her lines at a school play." "You see a woman commenting out loud about how fat another woman looks in her jeans." "You see a man quickly canceling a blind date as soon as he sees the woman." "You see a boy telling a woman that she looks just like her overweight bulldog."</p>
<p><u>Fairness</u></p>	<p>"You see a student copying a classmate's answer sheet on a makeup final exam.", "You see a runner taking a shortcut on the course during the marathon in order to win.", "You see a tenant bribing a landlord to be the first to get their apartment repainted.", "You see a soccer player pretending to be seriously fouled by an opposing player.", "You see someone cheating in a card game while playing with a group of strangers."</p>
<p><u>Authority</u></p>	<p>"You see a girl ignoring her father's orders by taking the car after her curfew.", "You see a woman refusing to stand when the judge walks into the courtroom.", "You see a girl repeatedly interrupting her teacher as he explains a new concept.", "You see a woman spray painting graffiti across the steps of the local courthouse.", "You see an intern disobeying an order to dress professionally and comb his hair."</p>
<p><u>Loyalty</u></p>	<p>"You see an employee joking with competitors about how bad his company did last year." "You see a coach celebrating with the opposing team's players who just won the game.", "You see a former US General saying publicly he would never buy any American product.", "You see a mayor saying that the neighboring town is a much better town.", "You see the US Ambassador joking in Great Britain about the stupidity of Americans."</p>
<p><u>Sanctity</u></p>	<p>"You see a man having sex with a frozen chicken before cooking it for dinner.", "You see a very drunk woman making out with multiple strangers on the city bus.", "You see a family eating the carcass of their pet dog that had been run over.", "You see a drunk elderly man offering to have oral sex with anyone in the bar.", "You see a teenager urinating in the wave pool at a crowded amusement park."</p>
<p><u>Liberty</u></p>	<p>"You see a man telling his fiance that she has to switch to his political party.", "You see a father requiring his son to become a commercial airline pilot like him.", "You see a man blocking his wife from leaving home or interacting with others.", "You see a manager coercing her employees into eating at her brother's diner.", "You see a man telling his girlfriend that she must convert to his religion."</p>
<p><u>Neutral</u></p>	<p>"You see someone using an old rotary phone and refusing to go buy a new one. ", "You see a man staying inside his home with the shades drawn on a rare sunny day. ", "You see a woman drinking her entire cup of coffee using a stirring spoon. ", "You see a woman eating dessert before her main entree arrives on the table. ", "You see a woman continuing to wear a large sun hat inside her apartment complex "</p>

Appendix E. Full List of Moral Foundation Vignettes and Average Foundation Ratings for 2022 Lucid Survey (Study 2)

[RATINGS RANGE FROM 1-5]	Care	Fair	Auth	Loyalty	Sanctity
Accidentally killing someone while drunk driving	4.67	3.54	2.68	2.12	4.11
Punching someone during an argument	3.83	2.61	2.31	2.55	3
Making fun of someone's clothes	3.58	2.27	1.46	2.18	2.43
Watching someone drown without trying to save them	4.33	3.36	2.25	3.29	3.96
Ignoring a homeless person asking for money	2.64	2.2	1.42	1.71	2.09
Hitting a misbehaving child	3.24	2.73	2.22	2.18	2.85
Putting your dog to sleep because its medical bills are too expensive	2.92	2.91	1.67	3.27	3.62
Calling someone ugly to their face	3.64	2.21	1.67	2.5	2.79
Laughing at someone with a disability	3.96	3.32	1.81	2.22	3.61
Using a racial or ethnic slur	3.82	3.26	2.43	2.82	3.66
Selling dangerous drugs to teenagers	4.36	3.88	3.25	2.92	4.09
Using a bonus to buy something nice for yourself instead of giving money to charity	1.63	1.7	1.12	1.66	1.33
Forgetting your best friend's birthday	2.33	1.89	1.28	3	1.87
Scolding a child for being too loud	2.19	1.53	1.11	1.83	1.63
Making fun of someone's foreign accent behind their back	3.12	3.27	1.96	2.21	2.54
Copying another person's answers during a test	2.86	3.81	2.87	2.38	2.89
Lying to the government about your income so you can pay less taxes	2.47	3.75	4.14	2.86	2.84
Bribing a government official	2.69	3.96	4.31	2.93	3.52
Robbing a bank	3.57	4.3	4.09	2.04	3.67
Using performance-enhancing drugs in a sport	2.68	4.19	3.21	2.75	3.3
Plagiarizing someone else's work	2.75	4.15	2.62	2.32	3.5
Having an affair with someone behind your spouse's back	4.33	4.05	2.25	4.5	4.14
Hiring a family member instead of the person who is most qualified	2.66	3.43	2.17	1.84	2.58
Selling someone a product that you know is defective	3.35	3.96	2.35	2.7	3.16
Embezzling money from the company you work for	3.31	4.19	4.44	3.85	3.73
Discriminating against someone because of their race or ethnicity	4.27	4.47	3.57	2.94	3.62
Cheating in a card game with friends to win extra money	2.74	4	2.23	3.93	3

	[RATINGS RANGE FROM 1-5]	Care	Fair	Auth	Loyalty	Sanctity
Voting more than once in an election		2.9	4.39	3.85	3.81	3.31
Lying about your experience in a job interview		2.15	3	2.61	2.06	2.72
Doing less work than everyone else on a team project		2.63	3.41	2.5	3.18	2.04
Switching the team you root for so you can be on the winning side		2	2.41	1.6	3.6	1.93
Spying on the United States for a foreign enemy		2.83	3.73	4.16	4.52	3.93
Gossiping about a close friend behind their back		3.33	2.64	1.68	3.66	2.83
Burning the American flag		2.78	2.38	3.21	3.38	3.58
Reporting a family member to the police for stealing		2.48	1.81	1.83	3.63	2.04
Buying a foreign instead of an American-made car		1.32	1.58	1.27	2	1.38
Faking a disability to get out of military service		3.07	4.08	3.75	3.48	3.71
Having sex with your best friend's spouse		4.33	4.26	2.57	4.43	4.27
Fighting for a foreign country in a war against the United States		2.86	2.87	3.64	3.88	3
Posting an embarrassing photo of your best friend on social media		3.47	2.88	1.55	3.92	2.68
Raising your kids in a foreign country		1.43	1.71	1.24	1.48	1.3
Ignoring a long-time friend to hang out with new people		2.83	2.91	1.36	3.4	2.21
Refusing to help a sibling who has fallen on hard times		3.48	2.42	1.73	3.88	3.03
Saying bad things about your childhood hometown		1.59	2	1.5	2.28	1.91
Betting against your own team and then intentionally losing the game		2.74	3.75	2.77	4.33	3.32
Refusing to stand for the President of the United States		1.72	2	3.52	2.43	2.52
Ignoring an email from your boss		2.04	1.87	3.56	2.66	1.92
Using someone's first name instead of their official title (e.g., "Dr.", "General")		1.54	1.81	2.56	1.5	1.96
Protesting against the government		1.69	1.68	3.26	2.27	1.67
Calling a police officer "pig"		2.77	2	3.5	2	2.85
Saying "fuck you" to your parents		3.48	2.48	4	3.56	3.3
Plotting to overthrow the government		3.11	3.19	4.73	3.9	3.59
Going on strike at work to demand higher wages		1.46	1.4	2.93	2.15	1.46
Disobeying a direct order from your commanding officer		1.96	2.96	4.5	4.14	3
Using a radar detector so you can speed without getting caught		1.88	2.49	2.96	1.57	2.06
Spraying graffiti on public buildings		2.25	2.64	3.17	2.37	3

[RATINGS RANGE FROM 1-5]	Care	Fair	Auth	Loyalty	Sanctity
Smoking marijuana	1.64	1.19	1.96	1.33	1.41
Wearing a t-shirt and shorts to a formal event	1.6	1.72	2.35	1.76	2.14
Participating in a riot	3.68	2.79	3.47	2.31	2.87
Laughing at your father in front of your friends	2.88	2.32	3.27	3.09	2.75
Having unprotected sex with multiple people in the same night	3.38	2.94	1.48	2.39	3.52
Defecating in a public park	3.14	2.38	2.88	1.83	3.64
Cursing God	2.55	2.22	3.44	2.78	3.69
Getting a tattoo on your face	1.48	1.48	1.52	1.28	1.79
Cooking and eating a rotting animal carcass that you found on the road	2.32	1.45	1.7	2	3.06
Not washing your hands after defecating	2.52	2.04	1.77	2.12	3.57
Eating insects	1.75	1.48	1.44	1	1.8
Drinking alcohol until you throw up	2.07	1.75	1.29	1.37	2.44
Throwing a holy book (e.g., Bible, Koran) in the trash	2.2	2.59	2.7	2.46	2.93
Having sex with a sibling	3.42	2.83	3.13	3.33	4.53
Performing sex acts for money	1.88	2.04	2.79	1.83	3.05
Pretending to be a monkey in front of a large group of people	1.64	1.39	1.52	1.38	1.88
Offering to sell your soul to someone else for \$1.00	2.63	2.03	2.25	2.31	3
Eating with your mouth open in front of other people	1.5	1.4	1.48	1.47	2
Throwing your deceased parents' ashes in a dumpster	3.19	2.93	3.2	4.23	4.15

Supplemental Appendix F. Alternative Measures of Ideology

The Dynata and Lucid samples in Study 1 and all data sets in Study 2 include policy items on various salient issues in American politics (see below). We use these items to construct three measures of issue-based left-right preferences. First, we extract factor scores from a principal factor model with a single latent dimension using all available policy items. We divide these scores into tertiles representing left-wing, moderate, and right-wing respondents. Recognizing the importance of distinguishing between the economic and social dimensions of ideology and, to be consistent with other work (Feldman and Johnston 2014), we also construct two measures tapping economic (e.g., minimum wage, importance of jobs versus environment) and social (e.g., abortion, gay marriage) issue preferences. We again extract factor scores from one-dimensional principal factors analyses, in each case limiting the items to only social or economic policy, and then dividing the scores into tertiles. Altogether, we have five distinct measures of left-right preferences: partisan identity, ideological identity, general issue preferences, economic issue preferences, and social issue preferences.

Rights for Sexual Minorities

Study 1

Which of the following comes closest to your own opinion regarding the legal status of sexual orientation in the United States?

- Gays and lesbians should receive the same legal protection from discrimination as racial, ethnic, and religious minorities
- Businesses and employers should NOT be punished by law for refusing service or employment based on sexual orientation

Study 2

Which of the following comes closest to your own opinion?

- Same-sex marriage should be legal
- Same-sex marriage should be illegal
- Same-sex marriage should be illegal, but same-sex couples should be allowed to form civil unions, giving them the same government benefits as marriage

Abortion

Studies 1 and 2

Which of the following comes closest to your own opinion regarding the legal status of abortion in the United States?

- By law, abortion should never be permitted
- The law should permit abortion only in cases of rape, incest, or when the woman’s life is in danger.
- The law should permit abortion for reasons other than rape, incest, or danger to the woman’s life, but only after the need for the abortion has been clearly established
- By law, a woman should always be able to obtain an abortion as a matter of personal choice

Immigration

Study 1

Which of the following comes closest to your own opinion regarding the number of immigrants from foreign countries who are permitted to come to the United States to live?

- The number of immigrants should be increased a lot.
- The number of immigrants should be increased a little.
- The number of immigrants should be decreased a little.
- The number of immigrants should be decreased a lot.

Affirmative Action/Aid to Blacks

Study 1

Some people say that because of past discrimination, blacks should be given preference in hiring and promotion. Others say that such preference in hiring and promotion of blacks is wrong because it gives blacks advantages they haven't earned. How much do you support or oppose preferential hiring of blacks?

- Support strongly
- Support somewhat
- Support a little
- Oppose a little
- Oppose somewhat
- Oppose strongly

Study 2

Which of the following comes closest to your own opinion?

- The government should do as much as possible to help the social and economic position of Blacks
- The government should do a little bit more to help the social and economic position of Blacks
- The government should do a little bit less to help the social and economic position of Blacks
- When it comes to their social and economic position, Blacks should help themselves

Taxes

Study 1

Which of the following is closest to your opinion regarding tax rates for wealthy Americans?

- Wealthy Americans should pay much higher taxes than they do now
- Wealthy Americans should pay somewhat higher taxes than they do now
- Wealthy Americans should pay somewhat lower taxes than they do now
- Wealthy Americans should pay much lower taxes than they do now

Insurance

Study 1

Some people feel there should be a government insurance plan which would cover all medical and hospital expenses for everyone. Suppose these people are at point 1 on the scale below. Others feel that all medical expenses should be paid by individuals through private insurance plans like Blue Cross or other company paid plans. Suppose these other people are at point 7. And of course, some people fall in-between at points 2, 3, 4, 5, and 6. What about you?

Minimum Wage

Study 1

The current federally mandated minimum wage is \$7.25 per hour. Which of the following comes closest to your own opinion?

- The federal minimum wage should be a lot higher.
- The federal minimum wage should be somewhat higher.
- The federal minimum wage should be a little higher.
- The federal minimum wage should be a little lower.
- The federal minimum wage should be somewhat lower.
- The federal minimum wage should be a lot lower.

Study 2

The current federally mandated minimum wage is \$7.25 an hour. Which comes closest to your own opinion?

- The federal minimum wage should be a lot higher
- The federal minimum wage should be a little higher
- The federal minimum wage should be a little lower
- The federal minimum wage should be a lot lower

Military spending

Study 1

Some people believe that we should greatly increase federal spending on the military, even if it means higher taxes, while others believe that we should greatly decrease spending on the military and use the savings for tax cuts or spending on social programs. Other people have opinions somewhere in-between. What about you?

- Increase spending on the military a great deal
- Increase spending on the military somewhat
- Increase spending on the military a little
- Decrease spending on the military a little
- Decrease spending on the military somewhat
- Decrease spending on the military a great deal

Environment versus Jobs

Study 2

Which of the following comes closest to your own opinion?

- The government should do as much as it can to regulate business in order to protect the environment and create jobs
- The government should do a little bit more to regulate business in order to protect the environment and create jobs
- The government should do a little bit less to regulate business in order to protect the environment and create jobs
- The government should not regulate business because it will not work and it will cost jobs

Supplemental Appendix G. Supplemental Results

Figure G1. Study 1 estimates using alternative ideology measures (available in Dynata Survey only)

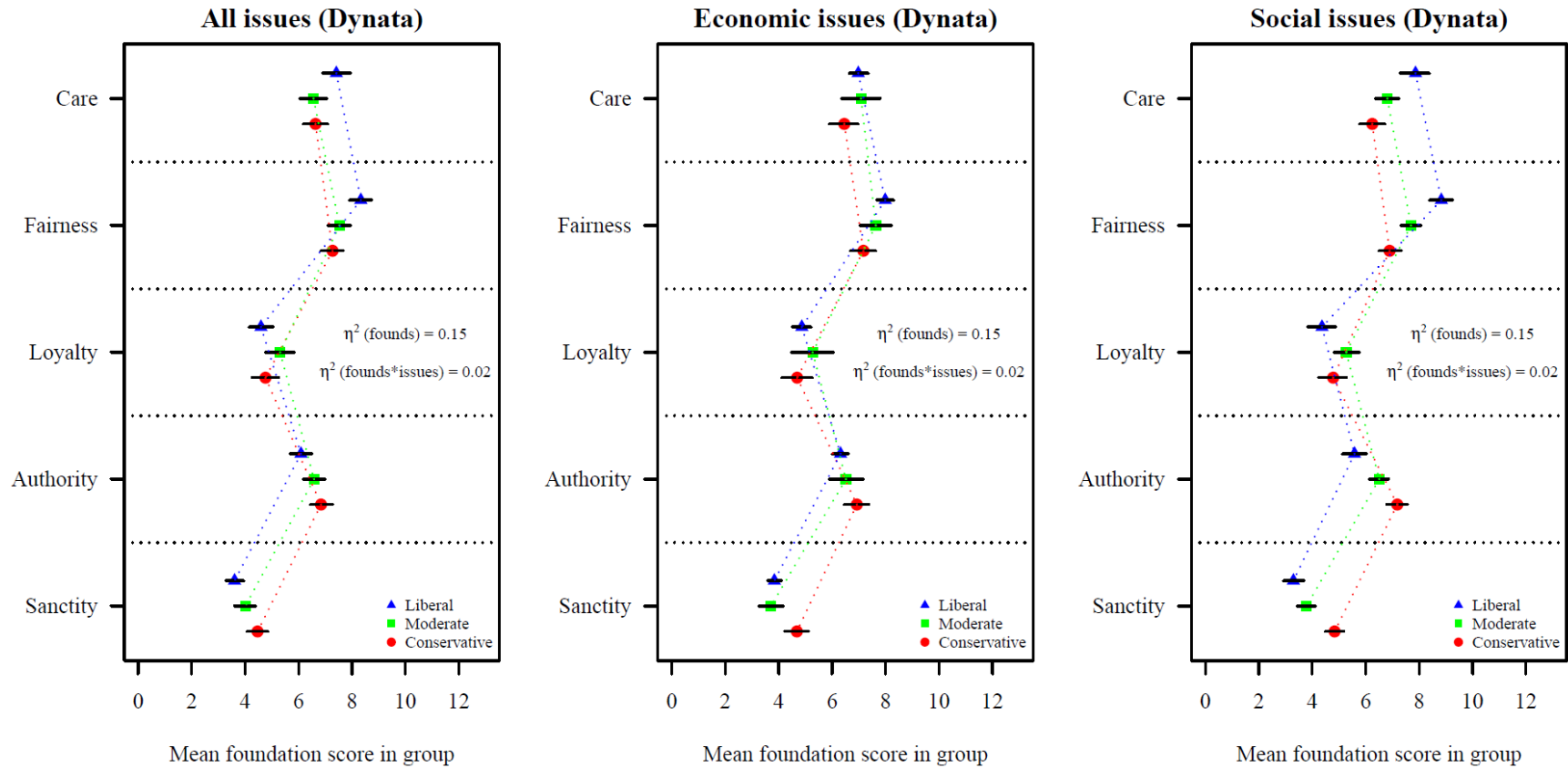


Figure G2. Study 2 estimates using alternative ideology measures

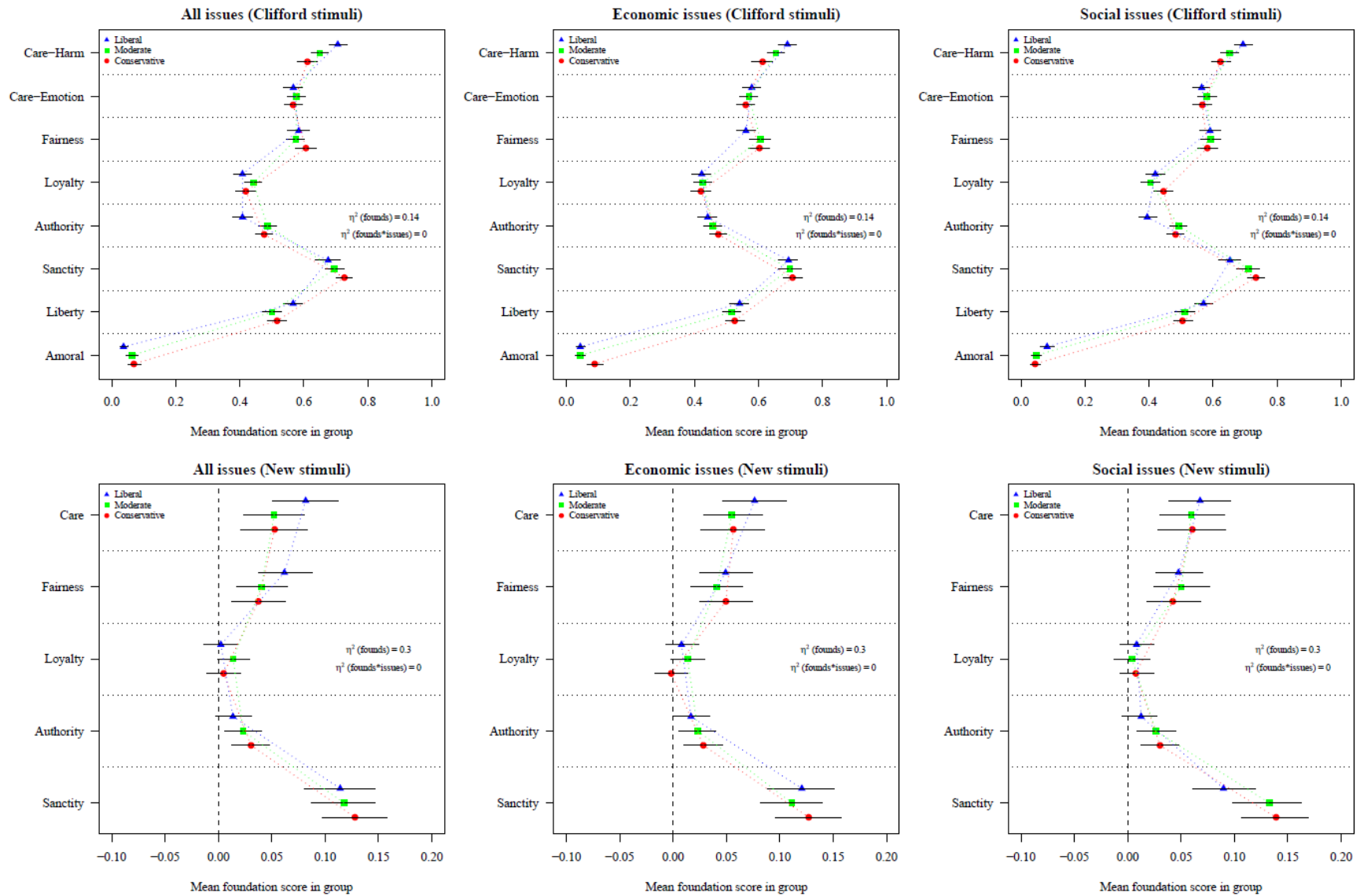


Figure G3. Estimates using the MFQ for partisanship and ideology

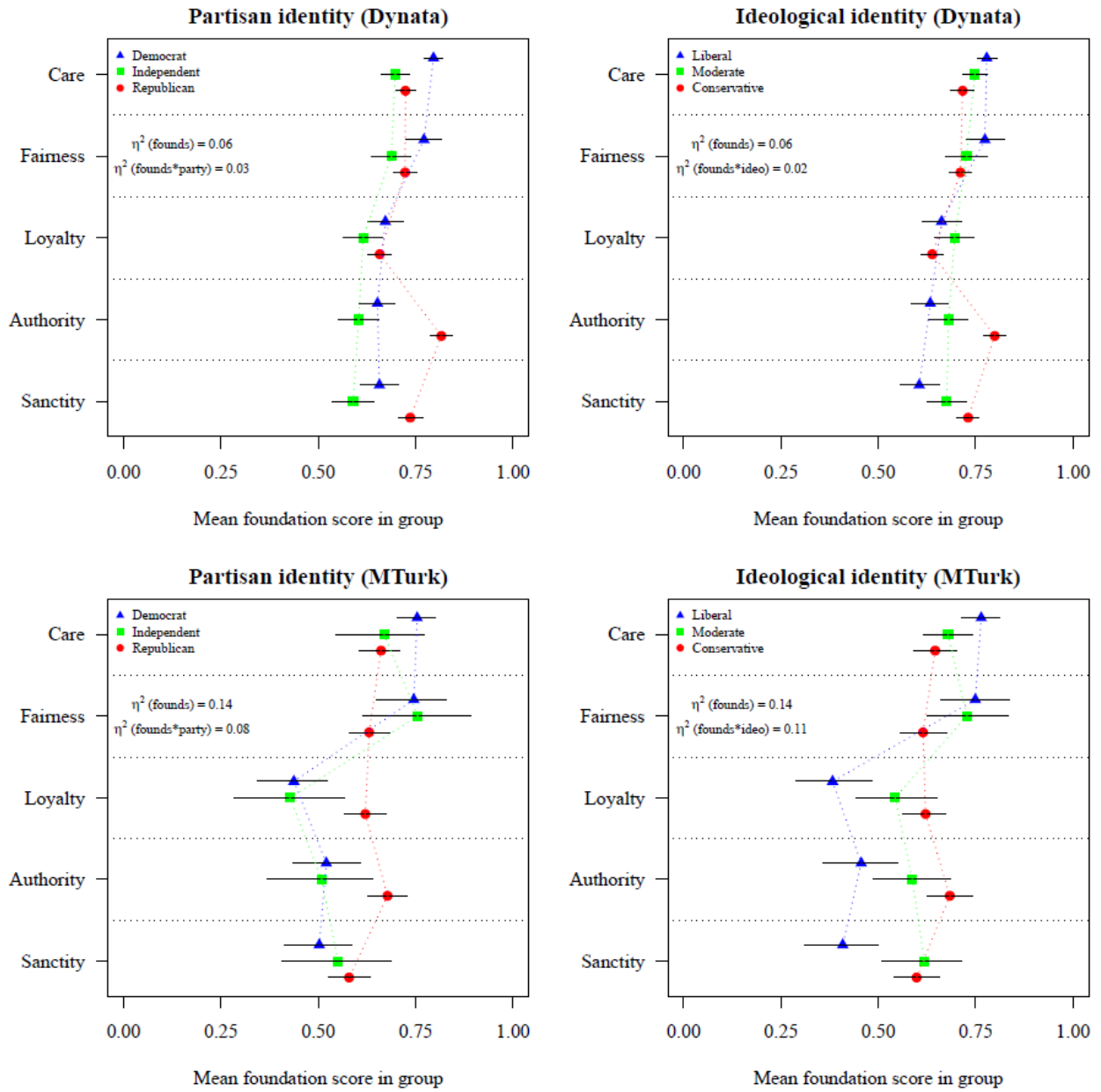
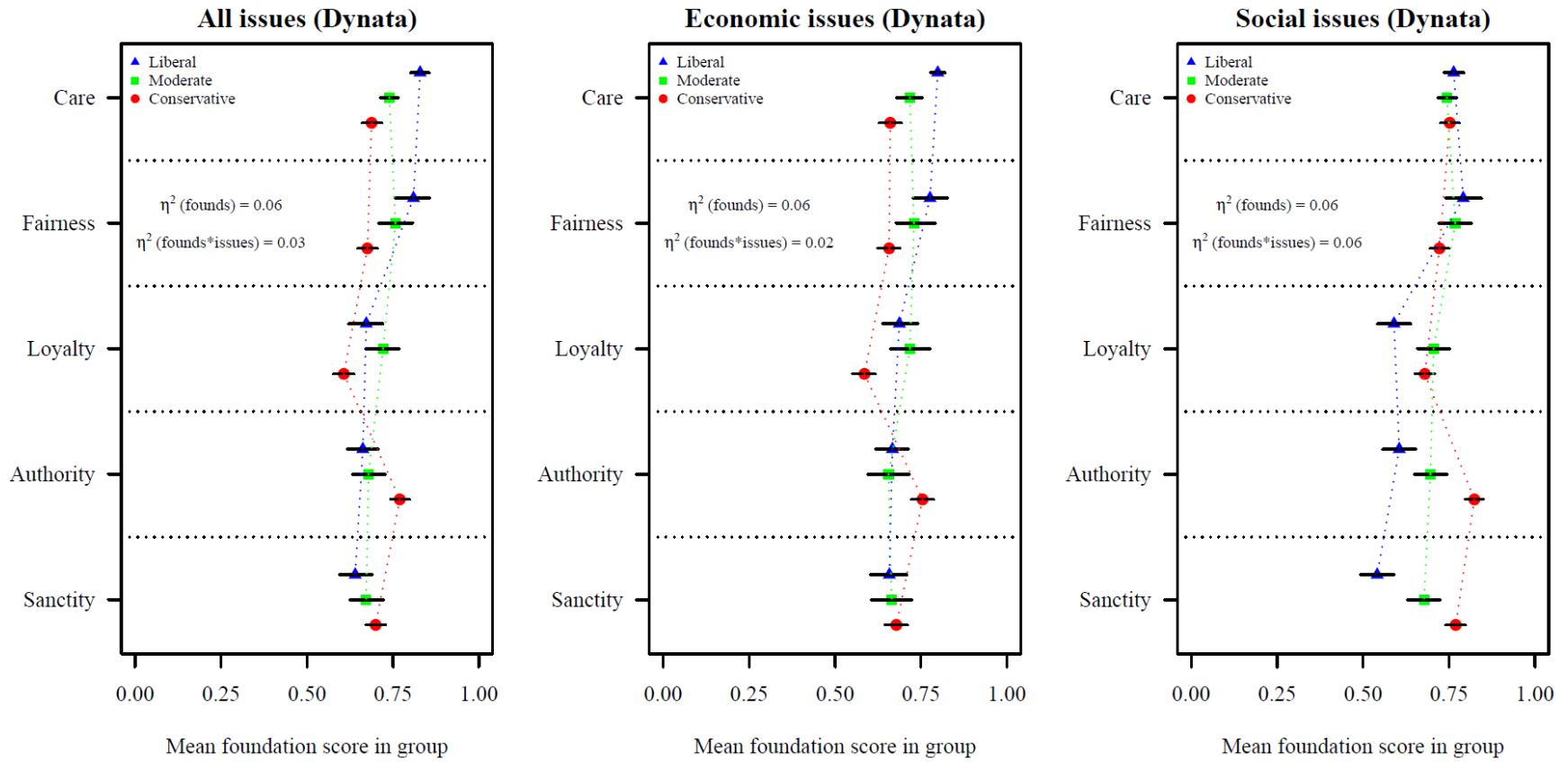


Figure G4. Estimates using the MFQ for alternative preference dimensions (Dynata only)



Supplemental Appendix H. References

- Alvarez, R Michael, Lonna Rae Atkeson, Ines Levin, and Yimeng Li. 2019. "Paying Attention to Inattentive Survey Respondents." *Political Analysis* 27 (2):145-162.
- Ansolabehere, Stephen, and Brian Schaffner. 2013a. CCES Common Content, 2012. edited by YouGov and Polimetrix: Harvard Dataverse.
- Ansolabehere, Stephen, and Brian Schaffner. 2013b. *cces_guide_2012.pdf*. In *CCES Common Content, 2012*, edited by YouGov and Polimetrix: Harvard Dataverse.
- Coppock, Alexander, and Oliver A McClellan. 2019. "Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents." *Research & Politics* 6 (1):2053168018822174.
- Davies, Caitlin L., Chris G. Sibley and James H. Liu. 2014. "Confirmatory Factor Analysis of the Moral Foundations Questionnaire: Independent Scale Validation in a New Zealand Sample." *Social Psychology* 45 (6): 431-436.
- Feldman, Stanley and Christopher Johnston. 2014. "Understanding the Determinants of Political Ideology: Implications of Structural Complexity." *Political Psychology* 35 (3): 337-358.
- Graham, Jesse, Brian A Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter H Ditto. 2011. "Mapping the Moral Domain." *Journal of Personality and Social Psychology* 101 (2):366.
- Iurino, Kathryn, and Gerard Saucier. 2018. "Testing Measurement Invariance of the Moral Foundations Questionnaire across 27 Countries." *Assessment*:1073191118817916.
- Nilsson, Artur, and Arvid Erlandsson. 2015. "The Moral Foundations Taxonomy: Structural Validity and Relation to Political Ideology in Sweden." *Personality and Individual Differences* 76:28-32.
- Yilmaz, Onurcan, Mehmet Harma, Hasan G. Bahçekapili and Sevim Cesur. 2016. "Validation of the Moral Foundations Questionnaire in Turkey and Its Relation to Cultral Schemas of Individualism and Collectivism." *Personality and Individual Differences* 99: 149-154.